



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number 177599

TO: Sheridan Swope
Art Unit: 1656
Location: rem/2B71/3C70
Serial Number: 09/940235

Friday, February 03, 2006

From: Beverly Shears
Location: Biotech-Chem Library
REM 1A54
Phone: 571-272-2528
beverly.shears@uspto.gov

Search Notes

Protein Sequence Searches – February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (uniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

Published Applications Database - November 2005

Published_Applications Nucleic Acid and Published_Applications Amino Acid database searches now generate two sets of results each. The Published_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published_Applications_New databases; older published applications make up the Published_Applications_Main databases.

Searches run against Nucleic Acid Published_Applications produce two sets of results, with the extensions **.rnpbm** (Published_Applications_NA_Main) and **.rnpbn** (Published_Applications_NA_New).
Searches run against Amino Acid Published_Applications produce two sets of results, with the extensions **.rapbm** (Published_Applications_AA_Main) and **.rapbn** (Published_Applications_AA_New).



From: Swope, Sheridan
Sent: *CRFE* Thursday, January 26, 2006 3:36 PM
To: STIC-Biotech/ChemLib
Subject: RE: 09/940,235

PLEASE ALSO ALIGN SID 9, 10, 11, AND 12.
Thanks!

-----Original Message-----

From: Swope, Sheridan
Sent: Wednesday, January 25, 2006 8:28 PM
To: STIC-Biotech/ChemLib
Subject: 09/940,235

For 09/940,235, pls search:

SID 9 against the NT and AA databases *NA 1541*

SID 10 against the NT and AA databases *NA 1661*

SID 11 against the NT and AA databases *NA 1782*

SID 12 against the NT and AA databases *NA 2096*

Sheridan Swope, Ph.D.
Patent Examiner, AU 1656
Recombinant Enzymes
571-272-0943 (voice)
E02B71 Remsen Bld (Office)
E03C70 Remsen Bld (Mailbox)

.....
Searcher: _____
Searcher Phone: _____
Date Searcher Picked up: _____
Date completed: _____
Searcher Prep Time: _____
Online Time: _____

.....
Type of Search
NA# _____ AA# _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure #: _____ Text: _____
Inventor: _____ Litigation: _____

.....
Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other (Specify): _____

.....
STIC
CM-1
Pre-2
Type of Search
N.A. Sequence
N.A. Sequence
Structure
Bibliographic

.....
IG
STN
Dialog
APS
Geninfo
SDC
DARC Questel
Other *CGW*

US-09-940-consensus	184	TCTCGATCCCGCGAAATTAAATACGACTACTATAGGAGACCAACACGGTTCCCTCTAGAG	tctcgatccgcgaaattaatacgactcactataggagaccacaacggtttccctctagag
US-09-940-consensus	1	TTTGTTHA	CTTTTAAAGAGGAGATATACCATGATaGCTGGtCCTGAaTGCTaCT
US-09-940-consensus	145	AATAATTTGT	TTTAACTTTAAGAGGAGATATACCATGATtGCTGACCTGAGTGCTGCT
US-09-940-consensus	195	AATAATTTGT	TTTAACTTTAAGAGGAGATATACCATGCTGCAAGCACACAGATTGTACC
US-09-940-consensus	245	AATAATTTGT	TTTAACTTTAAGAGGAGATATACCATGCTGCAAGCACACAGATTGTATACC
US-09-940-consensus		ataaatTTGT	TTTAACTTTAAGAGGAGATATACCATG-Tg--G-aC---Ag-----Ta-C-
US-09-940-consensus	56	AGA-CGTCCt	TCTGTaAAtAACAGCCAAATTGGTTGTtAGCGTTCTGTGTACTGTTGAGGGG
US-09-940-consensus	206	AGAcCGTCCa	TCTGTcAAcAACAGCCAAATTGGTTGTtAGCGTTCTGTGTACTGTTGAGGGG
US-09-940-consensus	256	CATAGCTGAGA	AGTGTtTTCATCATGCTCTGCGACTTCCTATGtGTGCGAGAAACCTGG
US-09-940-consensus	306	CATAGCTGAGA	AGTGTtTTCATCATGCTCTGCGACTTCCTATGtGTGCGAGAAACCTGG
US-09-940-consensus		--a--T--g	--t--t--A-----G-----T--C--TG--G-----G-----G--G--G
US-09-940-consensus	117	ACGAATCAAGACAT	TAGTCTTAAATTTTtGAAATCGATTAACATCAGACCTGCTCATG
US-09-940-consensus	267	ACGAATCAAGACAT	TAGTCTTAAATTTTtGAAATCGATTAACATCAGACCTGCTCATG
US-09-940-consensus	317	GAGAGGCGAGCGG	ACGATCACTTGCACTTCTAGAAATAGATGCAACGATCAGGACACAAG
US-09-940-consensus	367	GAGAGGCGAGCGG	ACGATCACTTGCACTTCTAGAAATAGATGCAACGATCAGGACACAAG
US-09-940-consensus		--GAA-----AG	-----TT--A-----CA-C-----C-CA-G
US-09-940-consensus	178	GAGGAAAGACAGAG	CGGCTTAAGTCCAAATCAAAACCAATTTGCTACTGATAGTGGCG
US-09-940-consensus	328	GAGGAAAGACAGAG	CGGCTTAAGTCCAAATCAAAACCAATTTGCTACTGATAGTGGCG
US-09-940-consensus	378	GACATCCTA	TAGAAATGGAGACACCTTGGAGCAAGAGATAATCGAGGAAACCTGCTCCAG
US-09-940-consensus	428	GACATCCTA	TAGAAATGGAGACACCTTGGAGCAAGAGATAATCGAGGAAACCTGCTCCAG
US-09-940-consensus		GA-----A-AGA	-----G-----A--T--A--AA-----A-T-G-----G--C--C--
US-09-940-consensus	239	GATGTCA	CATAAACTTGAGAAAGCTGACTTACTAAAGGCTATTCAAGACAAATGATCGCT
US-09-940-consensus	389	GATGTCA	CATAAACTTGAGAAAGCTGACTTACTAAAGGCTATTCAAGACAAATGATCGCT
US-09-940-consensus	439	TGCATCTGCA	CAGGCAACGCCGAGAGAGTGGAAAGTGTGAGAGGCACACCTCTGTGAGA
US-09-940-consensus	489	TGCATCTGCA	CAGGCAACGCCGAGAGAGTGGAAAGTGTGAGAGGCACACCTCTGTGAGA
US-09-940-consensus		-----TC-----A	-----G-----GA--G--A-----AA--G--A-----AC-----TG--C--C--
US-09-940-consensus	300	AACTGCCACAGTA	AAACGACGACTACTTTGAGGTCATTTGATTTTGGCAAGCGATGCAACCACTTA
US-09-940-consensus	450	AACTGCCACAGTA	AAACGACGACTACTTTGAGGTCATTTGATTTTGGCAAGCGATGCAACCACTTA
US-09-940-consensus	500	CCACATCGAGCGGAT	CTGGCCCTTCAACCGATGTTTCGTAATGCTGGACCTGAGTGGCTGCT
US-09-940-consensus	550	CCACATCGAGCGGAT	CTGGCCCTTCAACCGATGTTTCGTAATGCTGGACCTGAGTGGCTGCT
US-09-940-consensus		-----C-----A	-----G--C--T-----G-T--T-G-----TT-----C-----C-----
US-09-940-consensus	361	CTGATCGAAACGGCA	AGGTCTACTTTGCTGCAAAAGATGTTTCGGTAACTTGGCGACCCA

```
US-09-940- 511 CTGATCGAAGCGGCAAGGTCTACTTTGCTGACAAAGANGGTTGCGTAACTTGGCCGACCCA
US-09-940- 561 AGACCGTCCATCTGTCAACCAACAGCAATTTGGTTGTTAGCGTTGCTGGTACTGTTGAGGGG
US-09-940- 611 AGACCGTCACTGTCAACCAACAGCAATTTGGTTGTTAGCGTTGCTGGTACTGTTGAGGGG
consensus -----A-----T-----G--T-G-T-----TG--GA-----
US-09-940- 422 ACCTGTCCAAGAAATTTTCTAAGCGGCATGTCGGCTTAGACCATATAAGAAAAACCA
US-09-940- 572 ACCTGTCCAAGAAATTTTCTAAGCGGCATGTCGGCTTAGACCATATAAGAAAAACCA
US-09-940- 622 ACGAATCAAGCAATTTAGTCTTAAATTTTGAATTCGATCAACATCAGCACTGCTCATG
US-09-940- 672 ACGAATCAAGCAATTTAGTCTTAAATTTTGAATTCGATCAACATCAGCACTGCTCATG
consensus AC---TC-A---ATT--T--TAA-----T-----AC---A--A-----
US-09-940- 483 ATACAAAACCAAGCGAAATCTGTGATGGAATATACGTACAGTTTACTCCCTTAAACC
US-09-940- 633 ATACAAAACCAAGCGAAATCTGTGATGGAATATACGTACAGTTTACTCCCTTAAACC
US-09-940- 683 GAGGAAGACAGACAGCAAGGCTTAAAGTCCAAATCAAAACCAATTTGCTACTGATAGTGGGC
US-09-940- 733 GAGGAAGACAGACAGCAAGGCTTAAAGTCCAAATCAAAACCAATTTGCTACTGATAGTGGGC
consensus ----AAA--CA-----AA--CT--G-----A--A--A-----G-T-----T-----C
US-09-940- 544 CTGATGACGATTTTCAAGCCAGCTCTCAAGACTAAAGCTATTGAAAACTAGCTATCGG
US-09-940- 694 CTGATGACGATTTTCAAGCCAGCTCTCAAGACTAAAGCTATTGAAAACTAGCTATCGG
US-09-940- 744 GATGTCAATAAATTTGAGAAGCTGACTTACTAAAGGCTATTCAAGAACAAATTTGATCGCT
US-09-940- 794 GATGTCAATAAATTTGAGAAGCTGACTTACTAAAGGCTATTCAAGAACAAATTTGATCGCT
consensus -----T-AC-----C-----A-----CT-A--A--A-----A-GCTATT-AA-A-----
US-09-940- 605 TGACACCATCATCTCAAGAAATTTACTAGTCAAGCACAAAGCATTTTAAACAAAAACCAAC
US-09-940- 755 TGACACCATCATCTCAAGAAATTTACTAGTCAAGCACAAAGCATTTTAAACAAAAACCAAC
US-09-940- 805 AACGTCCACAGTAAACGACGACTACTTTGAGGTCAATTTGCAAGCGATGCAACCATTA
US-09-940- 855 AACGTCCACAGTAAACGACGACTACTTTGAGGTCAATTTGCAAGCGATGCAACCATTA
consensus -----CCA-----A-C-----T-----A-----A-----A-----
US-09-940- 666 CCAGGTATACGATTTTGAAGCGTACTCTCAATCGTCACTCATGACAAATGACATTTTCC
US-09-940- 816 CCAGGTATACGATTTTGAAGCGTACTCTCAATCGTCACTCATGACAAATGACATTTTCC
US-09-940- 866 CTGATCGAAACGGCAAGGCTCTACTTTGCTGCAAAAGATGGTTGCGTAACTTTGCCGACCCA
US-09-940- 916 CTGATCGAAACGGCAAGGCTCTACTTTGCTGCAAAAGATGGTTGCGTAACTTTGCCGACCCA
consensus C-----C-A-ACG-----T-----CAA---T-----TC---A---T---C-----C-----
US-09-940- 727 GTACGATTTTACCAATGGATCAAGAGTTTACTTTACCGTGTAAAAATCGGGAACAAGCTTTA
US-09-940- 877 GTACGATTTTACCAATGGATCAAGAGTTTACTTTACCGTGTAAAAATCGGGAACAAGCTTTA
US-09-940- 927 ACCTGTCCAAGAAATTTTGTCTAAGCGGCATGTCGGCTTAGACCATATAAGAAAAACCA
US-09-940- 977 ACCTGTCCAAGAAATTTTGTCTAAGCGGCATGTCGGCTTAGACCATATAAGAAAAACCA
consensus -----G-----T-----NAG-G--A--T-C-----T--A--AT-----A--AA-----A
```

```
US-09-940- 788 TAGGATCAATAAAAAATCTGCTCTGAATGAAGAAATAAACAAACACTGACCTGACTCTCTGAG
US-09-940- 938 TAGGATCAATAAAAAATCTGCTCTGAATGAAGAAATAAACAAACACTGACCTGACTCTGAG
US-09-940- 988 ATACAAAACCAAGCGAAATCTGTTGATGTGGAATATATCTGTACAGTTTACTTCCCTTAAACC
US-09-940- 1038 ATACAAAACCAAGCGAAATCTGTTGATGTGGAATATATCTGTACAGTTTACTTCCCTTAAACC
consensus -----A--A--AA---A--T---TGA-----A-ATA-----ACA-T---CT---T-----
US-09-940- 849 AAATATTACGTCCTTTAAAAAGGGAAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGA
US-09-940- 999 AAATATTACGTCCTTTAAAAAGGGAAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGA
US-09-940- 1049 CTGATGACGATTTTCAAGCCAGGCTCTCAAGATATCTAAAGTATTGAAAACTAGCTATCGG
US-09-940- 1099 CTGATGACGATTTTCAAGCCAGGCTCTCAAGATATCTAAAGTATTGAAAACTAGCTATCGG
consensus -----T-----A--A-G---AAG-----T-----A--C-C--C--T-G--G-
US-09-940- 910 AACTGTTACCATCAAAATACGTTGATGTCGATACCAACGAATTCGTTAAAGTGAGCAGCT
US-09-940- 1060 AACTGTTACCATCAAAATACGTTGATGTCGATACCAACGAATTCGTTAAAGTGAGCAGCT
US-09-940- 1110 TGACACCATCATCTCAAGAAATTTACTAGTCAAGCACAAAGCATTTTAAACAAAAACCAAC
US-09-940- 1160 TGACACCATCATCTCAAGAAATTTACTAGTCAAGCACAAAGCATTTTAAACAAAAACCAAC
consensus -----C-----A---T-A-----A---AC-AA-----T---AAA---A-----
US-09-940- 971 CTTAAACAGCTAGCGAAACGTAACCTTAGACTTTCAGAGATTTATACGATCCTCGTGATAAGGCT
US-09-940- 1121 CTTAAACAGCTAGCGAAACGTAACCTTAGACTTTCAGAGATTTATACGATCCTCGTGATAAGGCT
US-09-940- 1171 CAGGCTATACGATTTATGAACGTCGACTCTCTCAATCGTCACTCATGACAAATGACATTTTCC
US-09-940- 1221 CAGGCTATACGATTTATGAACGTCGACTCTCTCAATCGTCACTCATGACAAATGACATTTTCC
consensus C-----C-----G-----AAC-T-----A---T-A--C-----C-----C-
US-09-940- 1032 AAACTACTCTCAACCAATCTCGATGCTTTTGGTATTATGGAATATACCTTAACCTGGAAG
US-09-940- 1182 AAACTACTCTCAACCAATCTCGATGCTTTTGGTATTATGGAATATACCTTAACCTGGAAG
US-09-940- 1232 GTACGATTTTACCAATGGATCAAGATTTTACTTACCGTGTAAAAATCGGGAACAAGCTTTA
US-09-940- 1282 GTACGATTTTACCAATGGATCAAGATTTTACTTACCGTGTAAAAATCGGGAACAAGCTTTA
consensus --AC-A-T-TAC-A-----TC-A---TTT---TA-----TG-----A-A-C-----A-----
US-09-940- 1093 TAGAGGATAATACGATGACCAACCGTATCATAACCGTTTATATCGGGAAGCGACCCGA
US-09-940- 1243 TAGAGGATAATACGATGACCAACCGTATCATAACCGTTTATATCGGGAAGCGACCCGA
US-09-940- 1293 TAGAGTCAATAAAAAATCTGCTCTGAATGAAGAAATAAACCAACACTGACCTGATCTCTGAG
US-09-940- 1343 TAGAGTCAATAAAAAATCTGCTCTGAATGAAGAAATAAACCAACACTGACCTGATCTCTGAG
consensus TAG-----A--A--AT-----C-----A--A-----TG--C-----C-----
US-09-940- 1154 AGGAGAGAAATGCTAGCTACCATTTAGCTGTGTGGG-CAGCGCAaCagATtGfACcata
US-09-940- 1304 AGGAGAGAAATGCTAGCTATCATTTAGCCGCTGGTGGT-CAGCGCAGCAaATggtCagccc
US-09-940- 1354 AAATATTACGTCCTTTAAAAAGGGAAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGA
```

```
US-09-940- 1404 AATATTACGTCCTTAAATAAGGGGAAAGCCGTATGATCCCTTTGATCGCAGTCACCTTGA
consensus  A-----A---C-----Aa-A-----A-G--GT--g-----C---g--c--aGTC-cttga

US-09-940- 1215 gctgagaagtggtttTGatcaTgctgctggGaCtccTtaTgtGtCggaGaaAcgtgggAga
US-09-940- 1365 cAGTccccCGgtggCtGtCAGTcaaaAGcaAGcCcggTgTtATGaCaAtCgAaacaCtAtc
US-09-940- 1415 AACTGTTCCACCATCAAAATACGTTGATGTCGATACCAACGAATTGCTAAAAAAGTGACAGCT
US-09-940- 1465 AACTGTTCCACCATCAAAATACGTTGATGTCGATACCAACGAATTGCTAAAAAAGTGACAGCT
consensus  aactgttc-ccatc--atac-ttgatgtcGa-acca-c-aat-gCtaa-aA-tgagca-ct

US-09-940- 1276 AGccctAcCAAggcTGGatGatGgtagAttgtacttgccTggaGaaGcAGcgGacgcat
US-09-940- 1426 AGatAaAtCAAcagTGGaGcgGacCtAcctAGgTAAtgTgtGGTttGtActtGttatgg
US-09-940- 1476 CTTAAcAGCTAGCGAAACGTAACTTAGACTTCAGAGATTTATACGATCCCTCGTGATAAGGCT
US-09-940- 1526 CTTAACAGCTAGCGAAACGTAACTTAGACTTCAGAGATTTATACGATCCCTCGTGATAAGGCT
consensus  t--taacAgC-Agcg---gt-ac-taga-ttcaga-attt--acg-tcc-c-tga-aaggct

US-09-940- 1337 caecttGcacttcTagaAtatagatGcAacgAtCagGNaCaaggacatCtTatagaatttggA
US-09-940- 1487 AggaagCcgaggttttAACTGcgaaAGTaaCctGAAGCtgaagagacTTgtctTtGAcAAG
US-09-940- 1537 AAACTACTCTCAACAACATCTCGATGCTTTTGTGATTTATGACTATATACCTTAACTGGAAG
US-09-940- 1587 AAACTACTCTCAACAACATCTCGATGCTTTTGTGATTTATGACTATATACCTTAACTGGAAG
consensus  aaact-Ctctac-acaaA-ct-gatg-ttt-g-ta--at-gactatacCtTaaactggaaaaag

US-09-940- 1398 gACACctGGAGCAagaAggataaatcGaGaaAacctgctcCagtGcAtctgCaCaggcAAcG
US-09-940- 1548 TACACtGgGaaCACTtAcCgAgTgGtGacACtTAtgaCGTcctAaaGaCtCCatGAtCt
US-09-940- 1598 TAGAGGATAATCACGATGACACCAACCGTATCATAAcCGTTTATATGGCAAGCGACCCGA
US-09-940- 1648 TAGAGGATAATCACGATGACACCAACCGTATCATAAcCGTTTATATGGCAAGCGACCCGA
consensus  tA-A-ga--AtCAcga-gacaccaa-cGta-cataaccg-ttata-gggc-a-cgac-c-a

US-09-940- 1459 GccgagGagagTgGAagtGtGagaGGCaacacctctgTgcagaccACatcgagcgatctgg
US-09-940- 1609 GGGActGtAccTgCAtcgGgcTgGGcgaggagAatAagctGtACcaTcTaa
US-09-940- 1659 AGGAGAGAATGCTAGCTAtCATTTAGC ctAtgAtaaGatCgtTATAccgaagAag
US-09-940- 1709 AGGAGAGATGCTAGCTACCAATTTAGCtggTggTggccaggGcacaacgATtgtacccAta
consensus  -ggagaGaagt--a-ctagc-ttt-GC--g-g-tagtaa-gagcacc-tatac-ga-ca-g

US-09-940- 1520 ccccttcacgagatgttcgtagg
US-09-940- 1662
US-09-940- 1715 aacGAGAAGT
US-09-940- 1770 gctGAGAAGTgtttttgatcatgctgctgggaacttccatgtggtcgggaaaaacgtgggaga
consensus  -ccgagaagtg-t-t-----gctgctgggaacttccatgtggtcgggaaaaacgtgggaga

US-09-940- 1542
```

```
US-09-940- 1662
US-09-940- 1725 ttaCAGC
US-09-940- 1831 agccctaccaaggctggatgatggtagattgtacttgccttggggaaggCAGCGgaagcat
consensus
US-09-940- 1542 agccctaccaaggctggatgatggttagattgtacttgccttgggaga---cagcggagcat
US-09-940- 1662
US-09-940- 1732
US-09-940- 1892 cacttgcacttctagaaatagatgcaacgatcaggacacacaaggacatcctatagaattgga
consensus cacttgcacttctagaaatagatgcaacgatcaggacacacaaggacatcctatagaattgga
US-09-940- 1542
US-09-940- 1662
US-09-940- 1732 tACCTGC gtTatACAGG
US-09-940- 1953 gacacctgggcaagaaggataaatcgaggaaACCTGctccagtgcatcTgcACAGGcaacg
consensus gacacctgggcaagaaggataaatcgaggaa-acctgctccagtgca--t--acagggaacg
US-09-940- 1542
US-09-940- 1662
US-09-940- 1749 gACACCT ataccTgata
US-09-940- 2014 gccgaggagagtggaaagtgtgagaggcACACTctgtgcagaccacacatcgagcgaTctgg
consensus gccgaggagagtggaaagtgtgagagg-acacctctgtgcagaccacacatcga-----t-----

US-09-940- 1542
US-09-940- 1662
US-09-940- 1766 aCCCTaacgacAaaTaa
US-09-940- 2075 cCCCTtcaccgAtgTtcgtag
consensus -ccct-----a--t--gttag

Alignment score = -2900.00
Scoring matrix:
```

6	7	8	9
6	287	-1173	-2369
7		-423	-830
8			146
9			

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2006, 10:53:18 ; Search time 7834.07 Seconds
(without alignments)
11181.386 Million cell updates/sec

Title: US-09-940-235-9
Perfect score: 1541
Sequence: 1 ttgtttaacttaagaagg.....ccttcaccgatgttcgtag 1541

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues
Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_in.*
3: gb_env.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pr.*
9: gb_ro.*
10: gb_sta.*
11: gb_sy.*
12: gb_un.*
13: gb_vi.*
14: gb_htg.*
15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1134	73.6	2385	6	AR143998 Sequence
2	1132.6	73.5	1242	6	AR144000 Sequence
3	1132.6	73.5	1401	6	I05204 Sequence 5
4	1132.6	73.5	2568	1	STRSKC
5	1132.6	73.5	2568	6	A04926
6	1132.6	73.5	8931	1	SEDEXB
7	1127.8	73.2	2568	6	E00522
8	1120.2	72.7	7057	6	CQ979820
9	1118.6	72.6	2566	6	AR068768
10	1111.8	72.1	1245	1	S46536
11	1111.6	72.1	1257	6	A20015
12	1111.6	72.1	1257	6	I13203
13	1110.2	72.0	1407	6	E01413
14	1110.2	72.0	1512	6	A20016
15	1110.2	72.0	1512	6	I13204
16	1108.6	71.9	1317	6	A20009
17	1108.6	71.9	1317	6	I13197
18	1108.6	71.9	1335	6	A20006

19	1108.6	71.9	1335	6	I13194	I13194 Sequence 14
20	1108.6	71.9	1458	6	A20027	A20027 SEQ ID NO:
21	1108.6	71.9	1458	6	I13215	I13215 Sequence 42
22	1108.6	71.9	1467	6	A20030	A20030 SEQ ID NO:
23	1108.6	71.9	1467	6	I13218	I13218 Sequence 46
24	1108.6	71.9	2588	6	A20021	A20021 SEQ ID NO:
25	1108.6	71.9	2589	6	I13209	I13209 Sequence 34
26	1100.6	71.4	1473	1	SGSKG	XI3400 Streptococc
27	1099	71.3	1311	1	AY368335	AY368335 Streptoco
28	1084.4	70.4	1209	6	AR175891	AR175891 Sequence
29	1084.4	70.4	1209	6	AX030315	AX030315 Sequence
30	1084.4	70.4	1245	6	AR175892	AR175892 Sequence
31	1084.4	70.4	1245	6	AX030316	AX030316 Sequence
32	1083.8	70.3	1122	6	AR175893	AR175893 Sequence
33	1083.8	70.3	1122	6	AX030317	AX030317 Sequence
34	1083.8	70.3	1158	6	AR175894	AR175894 Sequence
35	1083.8	70.3	1158	6	AX030318	AX030318 Sequence
36	1078.8	70.0	2253	6	I13213	I13213 Sequence 39
37	1077.6	69.9	1119	6	I13206	I13206 Sequence 30
38	1075.8	69.8	2252	6	A20025	A20025 SEQ ID NO:
39	1065.6	69.1	1118	6	A20018	A20018 SEQ ID NO:
40	1066.2	65.3	1323	1	AY234129	AY234129 Streptoco
41	1066.2	65.3	1323	1	AY234137	AY234137 Streptoco
42	1066.2	65.3	50354	1	AE014169	AE014169 Streptoco
43	1066.2	65.3	110000	1	BA000034_16	Continuation (17 o
44	1066.2	65.3	110000	1	BA000034_17	Continuation (18 o
45	999.8	64.9	1323	1	AY234136	AY234136 Streptoco

ALIGNMENTS

RESULT 1
AR143998
LOCUS AR143998 2385 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 1 from patent US 6210667.
ACCESSION AR143998
VERSION AR143998.1 GI:15105865
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2385)
AUTHORS Reed G.L.
TITLE Bacterial fibrin-dependent plasminogen activator
JOURNAL Patent: US 6210667-A 1 03-APR-2001;
FEATURES Location/Qualifiers
source 1..2385
/organism="unknown"
/mol_type="unassigned DNA"

Query Match	73.6%	Score	1134	DB	6	Length	2385
Best Local Similarity	98.7%	Pred. No.	1.4e-278				
Matches	1143	Conservative	0	Mismatches	15	Indels	0
Gaps	0						
Qy	30	ATGATAGCTGGTCTCTGAATGGCTACTAGATCGTCTTCTGTAAATAACAGCAATTGGTT	89				
Db	1141	AGGATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCAATTAGTT	1200				
Qy	90	GTTAGCGTTCGTCTGTTGAGGGGACCAATCAAGACATTAGCTTTAAATTTTGGAA	149				
Db	1201	GTTAGCGTTCGTCTGTTGAGGGGACCAATCAAGACATTAGCTTTAAATTTTGGAA	1260				
Qy	150	ATCGATCTAATCAATCAGCACTGCTCTATGGAGGAAGACAGCAAGCGTTAAGTCCAAAA	209				
Db	1261	ATCGATCTAATCAATCAGCACTGCTCTATGGAGGAAGACAGCAAGCGTTAAGTCCAAAA	1320				
Qy	210	TCAAAACCAATTTGCTACTGTAGTGGCGCGATGTGCATATAAATTGAGAAGCTGACTTA	269				
Db	1321	TCAAAACCAATTTGCTACTGTAGTGGCGCGATGTGCATATAAATTGAGAAGCTGACTTA	1380				
Qy	270	CTAAAGGCTATTCAAGAACCAATTGATCGCTAAAGTCCACAGTACAGGACTACTTTTGAG	329				

Db	1381		CTAAAGCGTATTCAAGAACAAATGATCGCTAAACGTCACAGTACGACGACTACTTTTGAG	1440
Qy	330	GTCAATTGATTTTGCAGCGATGCAACCAATTA	CTGATCTGATAACGCAAGGCTCTACTTTTGCT	389
Db	1441	GTCAATTGATTTTGCAGCGATGCAACCAATTA	CTGATCTGATAACGCAAGGCTCTACTTTTGCT	1500
Qy	390	GACAAAGATGGTTTCGGTAAACCTTGC	CGACCCCAACCTGTCCAAAGAAATTTTTCGCTAAAGCGGA	449
Db	1501	GACAAAGATGGTTTCGGTAAACCTTGC	CGACCCCAACCTGTCCAAAGAAATTTTTCGCTAAAGCGGA	1560
Qy	450	CATGTGCGCGTTAGACCATATAAAGAAAAACCA	ATAAAAAACCAAGCGAAATCTGTGTGAT	509
.Db	1561	CATGTGCGCGTTAGACCATATAAAGAAAAACCA	ATAAAAAACCAAGCGAAATCTGTGTGAT	1620
Qy	510	GTGGAATATACTGTACAGTTTACTCCCTTAA	ACCCTGATGACGATTTTCAGACCAAGGCTCTC	569
Db	1621	GTGGAATATACTGTACAGTTTACTCCCTTAA	ACCCTGATGACGATTTTCAGACCAAGGCTCTC	1680
Qy	570	AAAGATACTAAGCTATTGAAACACTAGCTAT	CGGTGACCACTACATCTCAAGAAATTA	629
Db	1681	AAAGATACTAAGCTATTGAAACACTAGCTAT	CGGTGACCACTACATCTCAAGAAATTA	1740
Qy	630	CTAGCTCAAGCAAAAGCATTTTAAACAAAAAC	CACCCAGGCTATACGATTTATGAACGT	689
Db	1741	CTAGCTCAAGCAAAAGCATTTTAAACAAAAAC	CACCCAGGCTATACGATTTATGAACGT	1800
Qy	690	GACTCCTCAATCGTCACTCATGACAATGACA	ATTTTCCGTACGATTTTACCAATGGATCAA	749
Db	1801	GACTCCTCAATCGTCACTCATGACAATGACA	ATTTTCCGTACGATTTTACCAATGGATCAA	1860
Qy	750	GAGTTTACTTACCGTGTTTAAAAATCGGGACA	AGCTTTATAGGATCAATAAAAAATCTGGT	809
Db	1861	GAGTTTACTTACCGTGTTTAAAAATCGGGACA	AGCTTTATAGGATCAATAAAAAATCTGGT	1920
Qy	810	CTGAATCAAGAAATAAACAACACACTGACT	CTCTCGAGAAATATTACGTCCTTAAAAAA	869
Db	1921	CTGAATCAAGAAATAAACAACACACTGACT	CTCTCGAGAAATATTACGTCCTTAAAAAA	1980
Qy	870	GGGGAAAGCCGTATGATCCCTTTGATCGCAGT	CACTTGAAAATCTGTTACCATCAAAATAC	929
Db	1981	GGGGAAAGCCGTATGATCCCTTTGATCGCAGT	CACTTGAAAATCTGTTACCATCAAAATAC	2040
Qy	930	GTTGATGTCGATACCAACGAAATGCTCTAAA	AGTGAGCAGCTCTTTAAACAGCTAGCGAACGT	989
Db	2041	GTTGATGTCGATACCAACGAAATGCTCTAAA	AGTGAGCAGCTCTTTAAACAGCTAGCGAACGT	2100
Qy	990	AACCTTAGACTTCAGAGATTTTATACGATCC	TCTGTAAGGCTAAACTCTCTACACAAT	1049
Db	2101	AACCTTAGACTTCAGAGATTTTATACGATCC	TCTGTAAGGCTAAACTCTCTACACAAT	2160
Qy	1050	CTCGATGCTTTTGGTATTATGAGACTATAC	CTTAACTCGAAAAGTAGAGGATTAATCAGAT	1109
Db	2161	CTCGATGCTTTTGGTATTATGAGACTATAC	CTTAACTCGAAAAGTAGAGGATTAATCAGAT	2220
Qy	1110	GACACCAACCGTATCATAAACCGTTTATAT	GGGCAAGCGACCCGAAGGAGAGATGCTAGC	1169
Db	2221	GACACCAACCGTATCATAAACCGTTTATAT	GGGCAAGCGACCCGAAGGAGAGATGCTAGC	2280
Qy	1170	TACCAATTTAGCTGGTGGT	1187	
Db	2281	TATCAATTTAGCTATGAT	2298	

RESULT 2	1242 bp	DNA	linear	PAT 08-AUG-2001
AR144000				
LOCUS				
DEFINITION	Sequence 5 from patent US 6210667.			
ACCESSION	AR144000			
VERSION	AR144000.1			
KEYWORDS	GI:15105867			
SOURCE	Unknown.			
ORGANISM	Unknown.			

REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES	SOURCE	ORIGIN
-----------	---------	-------	---------	----------	--------	--------

REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
Unclassified.
1 (bases 1 to 1242)
Reed,G.L.
Bacterial fibrin-dependent plasminogen activator
Patent: US 6210667-A 5 03-APR-2001;
Location/Qualifiers
1. .1242

ORIGIN

Query Match	73.5%	Score 1132.6; DB 6; Length 1242;
Best Local Similarity	98.8%	Pred. No. 3.1e-278;
Matches 1141; Conservative	0;	Mismatches 14; Indels 0; Gaps 0;

Qy	33	ATAGCTGGTCTGTAATGGCTACTAGATCGTCTCTCTGTAATAACAGCGCAATTCGGTTGTT	92
Db	1	ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCGCAATTAGTTGTT	60
Qy	93	AGCGTTGCTGGTACTGCTTCAGCGGGACGAATCAAGACAATTAGTCTTTAAATTTTTTGAATC	152
Db	61	AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTTAAATTTTTTGAATC	120
Qy	153	GATCTAACATCACGACCTGCTCATCGAGGGAAGAAGACAGAGCAAGGCTTAAGTCCAAAATCA	212
Db	121	GATCTAACATCACGACCTGCTCATCGAGGGAAGAAGACAGAGCAAGGCTTAAGTCCAAAATCA	180
Qy	213	AAACCAATTTGCTACTGATAGTGGCGGATGTGCATATAAATTGAGAAAGCTGACTTACTA	272
Db	181	AAACCAATTTGCTACTGATAGTGGCGGATGTGCATATAAATTGAGAAAGCTGACTTACTA	240
Qy	273	AAGGCTATTCAAGAACAAATTGATCGCTAAAGCTGCACAGCTAACAGACGACTCTTTGAGGTC	332
Db	241	AAGGCTATTCAAGAACAAATTGATCGCTAAAGCTGCACAGCTAACAGACGACTCTTTGAGGTC	300
Qy	333	ATTGATTTTTCGAAGCGATGCAACCAATTACTGATTCGAAACGGCAAGGCTCTACTTTGCTGAC	392
Db	301	ATTGATTTTTCGAAGCGATGCAACCAATTACTGATTCGAAACGGCAAGGCTCTACTTTGCTGAC	360
Qy	393	AAAGATGGTTTCGGTAACTTTCGGGACCCAACTGTGTCGAAGAAATTTTTCGCTAAGCGGACAT	452
Db	361	AAAGATGGTTTCGGTAACTTTCGGGACCCAACTGTGTCGAAGAAATTTTTCGCTAAGCGGACAT	420
Qy	453	GTGCGGTTTAGAGCCATATAAGAGAAACCAATACAAACCAAGCGGAATCTGTTGATGTG	512
Db	421	GTGCGGTTTAGAGCCATATAAGAGAAACCAATACAAACCAAGCGGAATCTGTTGATGTG	480
Qy	513	GAATATACGTGACGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCAAGGCTCTCAAA	572
Db	481	GAATATACGTGACGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCAAGGCTCTCAAA	540
Qy	573	GATCTAAGGTTATTGAAAACACTAGCTATTCGGTGCACCAATCAATCTCAAGAAATTACTA	632
Db	541	GATCTAAGGTTATTGAAAACACTAGCTATTCGGTGCACCAATCAATCTCAAGAAATTACTA	600
Qy	633	GCTCAAGCAACAAGCAATTTTAAACAAAACCAACCCAGGCTATACGATTTATGAACGTTGAC	692
Db	601	GCTCAAGCAACAAGCAATTTTAAACAAAACCAACCCAGGCTATACGATTTATGAACGTTGAC	660
Qy	693	TCCTCAATCGTCACTCATGACAAATGACATTTTTCGGTACGATTTTACCAATGGATCAAGAG	752
Db	661	TCCTCAATCGTCACTCATGACAAATGACATTTTTCGGTACGATTTTACCAATGGATCAAGAG	720
Qy	753	TTTTACTACCGTTTAAAAATCGGGAACAGCTTTATAGGATCAATAAAAAATCTGGTCTG	812
Db	721	TTTTACTACCGTTTAAAAATCGGGAACAGCTTTATAGGATCAATAAAAAATCTGGTCTG	780
Qy	813	AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAGTCTCTTAAAAAAGGG	872
Db	781	AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAGTCTCTTAAAAAAGGG	840
Qy	873	GAAGAAGCGGTATGATCCCTTTGATCGCAGTCACTTTGAAACCTGTTCAACCATCAATACGTT	932

Db 841 GAAAGCCGATGATCCCTTTGATCGCAGTCACCTTGAAACTGTTCCACATCAATACGTT 900
QY 933 GATGTCGATACCAACGAATGCTAAAGATGAGCAGCTCTTAAACAGCTAGCGAAGCTAAC 992
Db 901 GATGTCGATACCAACGAATGCTAAAGATGAGCAGCTCTTAAACAGCTAGCGAAGCTAAC 960
QY 993 TTAGACTTCAGAGATTTATACGATCTCTGATGATAAGGCTAAACTACTCTACAACAATCTC 1052
Db 961 TTAGACTTCAGAGATTTATACGATCTCTGATGATAAGGCTAAACTACTCTACAACAATCTC 1020
QY 1053 GATGCTTTTGGTATTATGAGTACTACCTTAACCTGGAAGATGAGGATAATCAGATGAC 1112
Db 1021 GATGCTTTTGGTATTATGAGTACTACCTTAACCTGGAAGATGAGGATAATCAGATGAC 1080
QY 1113 ACCAACCGTATCAATACCGTTTATATGCGGCAAGCGACCCGAAGGAGAGAAATGCTAGCTAC 1172
Db 1081 ACCAACCGTATCAATACCGTTTATATGCGGCAAGCGACCCGAAGGAGAGAAATGCTAGCTAT 1140
QY 1173 CATTTAGCTGGTGGT 1187
Db 1141 CATTTAGCTATGAT 1155

RESULT 3
I05204
LOCUS 1401 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 5 from Patent EP 0248227.
ACCESSION I05204
VERSION I05204.1 GI:591209
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1401)
AUTHORS Hagenson,M.J. and Stroman,D.W.
TITLE Yeast production of streptokinase
JOURNAL Patent: EP 0248227-A1 5 09-DEC-1987;
FEATURES
source 1. 1401
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 73.5%; Score 1132.6; DB 6; Length 1401;
Best Local Similarity 98.4%; Pred. No. 3.1e-278;
Matches 1144; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
QY 25 ATACCATGATAGTGGTCTCTGATGGCTACTAGATCGTCTCTCTGTAAATAACAGCCAAAT 84
Db 2 ATTCCATGATGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCCAAAT 61
QY 85 TGGTTGTTAGCTGTTGTTAGTGGGACGAATCAAGACATTAGTCTTAAATTTT 144
Db 62 TAGTTGTTAGCTGTTGTTAGTGGGACGAATCAAGACATTAGTCTTAAATTTT 121
QY 145 TTGGAATCCATCTAATCAACGACCTGCTCATGGAGGAAGACAGACGACGCTTAAGTC 204
Db 122 TTGGAATCCATCTAATCAACGACCTGCTCATGGAGGAAGACAGACGACGCTTAAGTC 181
QY 205 CAAATCAAAACCAATTTGCTACTGATAGTGGCCGCGATGTCACATAAACTTGAGAAAGCTG 264
Db 182 CAAATCAAAACCAATTTGCTACTGATAGTGGCCGCGATGTCACATAAACTTGAGAAAGCTG 241
QY 265 ACTTACTAAAGGCTATTCAAGAACAAATTGATCGCTAACGTCCTCAAGTAACGACGACTACT 324
Db 242 ACTTACTAAAGGCTATTCAAGAACAAATTGATCGCTAACGTCCTCAAGTAACGACGACTACT 301
QY 325 TTGAGGTCAATGATTGTCAGGCGATGCAACCAATTAATGATCGAAGCGGCAAGGCTTACT 384
Db 302 TTGAGGTCAATGATTGTCAGGCGATGCAACCAATTAATGATCGAAGCGGCAAGGCTTACT 361
QY 385 TTGCTGACAAAGATGTTGCTGCTAAACCTTGCAGCACCAACCTGTCCAGGAATTTTCTTAA 444

Db 362 TTGCTGCAAAAGATGGTTTCGGTAACTTTGCCGACCCCAACCTGTCCAAGAAATTTTTCGTAA 421
QY 445 GCGGACATGTGCGGCTTTAGACCATATAAGAGAAAAACCAATACAAACCAAGCGAAATCTG 504
Db 422 GCGGACATGTGCGGCTTTAGACCATATAAGAGAAAAACCAATACAAACCAAGCGAAATCTG 481
QY 505 TTGATGTGGAATATACCTGTACAGTTTACTCCCTTTAAACCCCTGATGACGATTTTCAGACAG 564
Db 482 TTGATGTGGAATATACCTGTACAGTTTACTCCCTTTAAACCCCTGATGACGATTTTCAGACAG 541
QY 565 GTCTCAAGATATACCTGTGAAACATAGCTATCGGTATCGGTGACACCAATCACTCTCAAG 624
Db 542 GTCTCAAGATATACCTGTGAAACATAGCTATCGGTATCGGTGACACCAATCACTCTCAAG 601
QY 625 AATTACTAGCTCAAGACCAAGCAATTTTAAACAAAAACCCAGGCTTATACGATTTTATG 684
Db 602 AATTACTAGCTCAAGACCAAGCAATTTTAAACAAAAACCCAGGCTTATACGATTTTATG 661
QY 685 AACGTGACTCTCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGG 744
Db 662 AACGTGACTCTCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGG 721
QY 745 ATCAAGAGTTTACTTACCGTGTAAATAATCGGGAACAAAGCTTATAGGATCAATAAAAAT 804
Db 722 ATCAAGAGTTTACTTACCGTGTAAATAATCGGGAACAAAGCTTATAGGATCAATAAAAAT 781
QY 805 CTGCTCTGAATGAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTTACGTCCTTA 864
Db 782 CTGCTCTGAATGAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTTACGTCCTTA 841
QY 865 AAAAAGGGGAAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTTACCAATCA 924
Db 842 AAAAAGGGGAAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTTACCAATCA 901
QY 925 AATACTGATGATGATACCAACGAATGCTTAAAGTGAAGCAGCTCTTAAACAGCTAGCG 984
Db 902 AATACTGATGATGATACCAACGAATGCTTAAAGTGAAGCAGCTCTTAAACAGCTAGCG 961
QY 985 AACGTAACTTAGACTTCAGAGATTTATACGATCTCTGATGATGATGATGATGATGATGATGAT 1044
Db 962 AACGTAACTTAGACTTCAGAGATTTATACGATCTCTGATGATGATGATGATGATGATGATGAT 1021
QY 1045 ACAATCTCGATGCTTTTGGTATTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1104
Db 1022 ACAATCTCGATGCTTTTGGTATTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1081
QY 1105 ACGATGACACCAACCGCTATCAATACCGTTTATATGCGCAAGCGACCCGAAAGGAGAGATG 1164
Db 1082 ACGATGACACCAACCGCTATCAATACCGTTTATATGCGCAAGCGACCCGAAAGGAGAGATG 1141
QY 1165 CTAGCTACCAATTTAGCTGGTGGT 1187
Db 1142 TCAGCTATCATTTAGCTATGAT 1164

RESULT 4
STRSC
LOCUS 2568 bp DNA linear BCT 26-APR-1993
DEFINITION Streptococcus equisimilis (H46A) streptokinase gene, complete cds.
ACCESSION K02986
VERSION K02986.1 GI:153808
KEYWORDS streptokinase.
SOURCE Streptococcus dysgalactiae subsp. equisimilis
ORGANISM Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
REFERENCE 1 (bases 1 to 2568)
AUTHORS Malke,H., Roe,B. and Ferretti,J.J.
TITLE Nucleotide sequence of the streptokinase gene from Streptococcus equisimilis H46A
JOURNAL Gene 34 (2-3), 357-362 (1985)
PUBMED 2989113
COMMENT Original source text: S.equisimilis (strain H46A) DNA, clone pMP5.

Draft entry and hard copy of sequence for [1] kindly provided by J.J.Ferretti, 03-SEP-1985.
The -35 and -10 regions are located at positions 760-765 and 781-786 respectively and an SD sequence at 808-813. Downstream from the coding region inverted repeats (positions 2176-2190 and 2203-2217) are thought to function as transcription terminators. The nucleotide sequence of skc does not support the hypothesis that the gene has evolved by duplication and fusion, as suggested by internal two-fold AA homologies of its product.

FEATURES

Location/Qualifiers
1. .2568
/organism="Streptococcus dysgalactiae subsp. equisimilis"
/mol_type="genomic DNA"
/db_xref="taxon:119602"
794_ .>2141
/product="skc mRNA"
819. .2141
/note="prestreptokinase"
/codon_start=1
/transl_table=11
/protein_id="AAA26974.1"
/db_xref="GI:153809"
/translation="MKNYLSFGMFALLFALFTGVNVSQAIAGPEWLLDRPSVNNSQL
VVSAGTVGTNQDLSLKEFDLTSRPAHGKTEQGLSPKSPFADSGAMSHKLEK
ADLLKAIQRLIANVHNSDDYFVIDFASDATTDRNGKVFADKQSVLTPQVQE
FLLSGHVRVPYKEKPIQNOAKSVDEVTVQTPFLNPDDDFRPLKDTLLKTLAIGD
TITSQELLAQAOILKNHRYGTYIYERDSSIVTHDNDIFRTILPMDQBYRYRVNREQ
AYRNKKSGLNEINNDLISEKYIVLKKEKPYDFPDRSHLKLFTIKIVDVVDNELL
KSEQLLTASERLDFDLDPDRKAKLLYNLDLDAFGIMDYTLTGKVEDNHDNTRIIT
VYMKRPEGENASVYHLAYDKDRYTEEREVYSYLYRTGTPIDPNPDK"
819. .896
/note="streptokinase signal peptide"
mat_peptide
897. .2138
/product="streptokinase"
ORIGIN 5 bp upstream of PstI site.
Query Match 73.5%; Score 1132.6; DB 1; Length 2568;
Best Local Similarity 98.8%; Pred. No. 3.1e-278;
Matches 1141; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 33 ATAGCTGGTCTGAATGGCTACTAGATCGCTCTTCTGTAATAACAGCCAAATGGTGT 92
DB 897 ATTGCTGGACCTGAGTGGCTGTAGACCGCTCCATCTGTCAACACAGCCAAATGATGTT 956
QY 93 ACCGTGTGCTGCTACTGTTGAGGGGACGAATCAAGACATAGTCTTAAATTTTGAATC 152
DB 957 ACCGTGTGCTGCTACTGTTGAGGGGACGAATCAAGACATAGTCTTAAATTTTGAATC 1016
QY 153 GATCTAAACATCAGACCTGCTCATGGAGGAAGACAGACAGCGCTTAAGTCCAAATCA 212
DB 1017 GATCTAAACATCAGACCTGCTCATGGAGGAAGACAGACAGCGCTTAAGTCCAAATCA 1076
QY 213 AAACCATTTGCTACTGATAGTGGCGGATGTCTACATAAACTTGAGAAAGCTGACTACTA 272
DB 1077 AAACCATTTGCTACTGATAGTGGCGGATGTCTACATAAACTTGAGAAAGCTGACTACTA 1136
QY 273 AAGGCTATTCAAGAACAAATTTGATCGCTAACGCTCCACAGTAACGACGACTTTGAGGTC 332
DB 1137 AAGGCTATTCAAGAACAAATTTGATCGCTAACGCTCCACAGTAACGACGACTTTGAGGTC 1196
QY 333 ATTGATTTTGAAGCGCATCAACATTTACTGATCGAAACGGCAGGCTACTTTGCTGAC 392
DB 1197 ATTGATTTTGAAGCGCATCAACATTTACTGATCGAAACGGCAGGCTACTTTGCTGAC 1256
QY 393 AAGATGTTTGGTAACTTGGCGACCCCAACCTGTCCAAAGAAATTTTGTCTAAGCGGACAT 452
DB 1257 AAGATGTTTGGTAACTTGGCGACCCCAACCTGTCCAAAGAAATTTTGTCTAAGCGGACAT 1316
QY 453 GTGCGGTTAGACCATATATAAGAAAAACAATACAAAAACAAGCGAAATCTGTTGATGTG 512
DB 1317 GTGCGGTTAGACCATATATAAGAAAAACAATACAAAAACAAGCGAAATCTGTTGATGTG 1376
QY 513 GAATATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCGAGTCTCAA 572

Db 1377 GAATATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCGAGTCTCAA 1436
QY 573 GATACTAAGCTATTGAACAACACTAGCTATCGGTGACACCATCATCATCTCAAGATTTACTA 632
Db 1437 GATACTAAGCTATTGAACAACACTAGCTATCGGTGACACCATCATCATCTCAAGATTTACTA 1496
QY 633 GCTCAAGCACAAGAGCATTTTAAACAAACACACCCAGGCTATACGATTTATGAACGTCAC 692
Db 1497 GCTCAAGCACAAGAGCATTTTAAACAAACACACCCAGGCTATACGATTTATGAACGTCAC 1556
QY 693 TCCTCAATCGTCTACTCATGACAATGACATTTTCGTCAGATTTTACCAATGGATCAAGAG 752
Db 1557 TCCTCAATCGTCTACTCATGACAATGACATTTTCGTCAGATTTTACCAATGGATCAAGAG 1616
QY 753 TTTTACTACCGTGTAAATAATCGGAAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 812
Db 1617 TTTTACTACCGTGTAAATAATCGGAAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 1676
QY 813 AATGAAGAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 872
Db 1677 AATGAAGAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 1736
QY 873 GAAAGCCGTATGATCCCTTTTGTATCGCAGTCACTTGAAACTGTTTCAACATCAATACGTT 932
Db 1737 GAAAGCCGTATGATCCCTTTTGTATCGCAGTCACTTGAAACTGTTTCAACATCAATACGTT 1796
QY 933 GATGTGATACCAACGAATTTGCTAAAGTGAGCAGCTCTTAAACAGCTAGGCAACGTAAC 992
Db 1797 GATGTGATACCAACGAATTTGCTAAAGTGAGCAGCTCTTAAACAGCTAGGCAACGTAAC 1856
QY 993 TTAGACTTCAGAGATTTATACGATCTCTCGTGATAGGCTAAACTCTCTACACAATCTC 1052
Db 1857 TTAGACTTCAGAGATTTATACGATCTCTCGTGATAGGCTAAACTCTCTACACAATCTC 1916
QY 1053 GATGCTTTTGGTATTATGAGTATACCTTAACCTGGAAGAGTAGAGGATAATCAGATGAC 1112
Db 1917 GATGCTTTTGGTATTATGAGTATACCTTAACCTGGAAGAGTAGAGGATAATCAGATGAC 1976
QY 1113 ACCAACCGTATACATACCGTTTATATGGCAAGCGACCCGAGGAGAGAGATGCTAGTAC 1172
Db 1977 ACCAACCGTATACATACCGTTTATATGGCAAGCGACCCGAGGAGAGAGATGCTAGTAC 2036
QY 1173 CATTTAGCTGGTGGT 1187
Db 2037 CATTTAGCTGGTGGT 2051
RESULT 5
A04926
LOCUS S.equisimilis skc gene for streptokinase. 2568 bp DNA linear PAT 15-JUL-1993
DEFINITION A04926
ACCESSION A04926
VERSION A04926.1 GI:412219
KEYWORDS streptokinase.
SOURCE Streptococcus dysgalactiae subsp. equisimilis
ORGANISM Streptococcus dysgalactiae subsp. equisimilis
REFERENCE 1 (bases 1 to 2568)
AUTHORS Ferretti J.J. and Malke H.
TITLE Streptokinase-coding recombinant vectors
JOURNAL Patent: EP 015137-A 1 14-AUG-1985;
THE BOARD OF REGENTS OF THE UNIVERSITY OF OKLAHOMA
FEATURES
1. .2568
Location/Qualifiers
/organism="Streptococcus dysgalactiae subsp. equisimilis"
/mol_type="unassigned DNA"
/sub_species="equisimilis"
/db_xref="taxon:119602"
ORIGIN
Query Match 73.5%; Score 1132.6; DB 6; Length 2568;

Best Local Similarity 98.8%; Pred. No. 3.1e-278; Matches 1141; Conservative 0; Mismatches 14; Indels 0; Gaps 0;			
Qy	33	ATAGCTGGTCCCTGAATGGCTACTAGATCGTCCTTCTGTAATAATACAGCCAAATGGTTGTT	92
Db	897	ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAAATAGTTGTT	956
Qy	93	ACGGTGTGCTGTACTGTTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTTTGAATC	152
Db	957	ACGGTGTGCTGTACTGTTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTTTGAATC	1016
Qy	153	GATCTAATCATCAGCACTCTCATGTGAGGAAAGACAGACAGCGCTTAAGTCCAAATCA	212
Db	1017	GATCTAATCATCAGCACTCTCATGTGAGGAAAGACAGACAGCGCTTAAGTCCAAATCA	1076
Qy	213	AAACCAATTTGCTACTCATAGTGGCGGATGTACATAAATCTCAGAAAGCTGACTACTA	272
Db	1077	AAACCAATTTGCTACTCATAGTGGCGGATGTACATAAATCTCAGAAAGCTGACTACTA	1136
Qy	273	AAGGCTATTCAAGAAACAATTTGATCGCTAACGCTCCACAGTAACGACGACTACTTTGAGTC	332
Db	1137	AAGGCTATTCAAGAAACAATTTGATCGCTAACGCTCCACAGTAACGACGACTACTTTGAGTC	1196
Qy	333	ATTGATTTTGAAGCGATGCAACCATATCTGATCGAAAAACGGCAAGGTCTACTTTGCTGAC	392
Db	1197	ATTGATTTTGAAGCGATGCAACCATATCTGATCGAAAAACGGCAAGGTCTACTTTGCTGAC	1256
Qy	393	AAAGATGGTTCGGTACCTTCGCGCCACCTGCTCCAGAAATTTTGTCTAAGCGGCAT	452
Db	1257	AAAGATGGTTCGGTACCTTCGCGCCACCTGCTCCAGAAATTTTGTCTAAGCGGCAT	1316
Qy	453	GTGCGGTTAGACCATATATAAGAAAAACCAATACAAAAACCAAGCAAACTGTTGATGTG	512
Db	1317	GTGCGGTTAGACCATATATAAGAAAAACCAATACAAAAACCAAGCAAACTGTTGATGTG	1376
Qy	513	GAATATCTGTACAGTTTACTCCCTTAAACCCCTGTATGACGANTTCAGACAGGCTCTCAA	572
Db	1377	GAATATCTGTACAGTTTACTCCCTTAAACCCCTGTATGACGANTTCAGACAGGCTCTCAA	1436
Qy	573	GATCTAAGCTATTGAAACACTAGTATCGGTGACACCATCTCAAGTAATTA	632
Db	1437	GATCTAAGCTATTGAAACACTAGTATCGGTGACACCATCTCAAGTAATTA	1496
Qy	633	GCTCAAGCAACAAGCATTTTAAACAAAAACCAACCCAGGCTATACGATTTATGACGTGAC	692
Db	1497	GCTCAAGCAACAAGCATTTTAAACAAAAACCAACCCAGGCTATACGATTTATGACGTGAC	1556
Qy	693	TCCTCAATCGTCACTCATGCAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG	752
Db	1557	TCCTCAATCGTCACTCATGCAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG	1616
Qy	753	TTTACTTACCGTGTAAATAATCGGGAACAGCTTATAGGATCAATAAAAAATCTGCTG	812
Db	1617	TTTACTTACCGTGTAAATAATCGGGAACAGCTTATAGGATCAATAAAAAATCTGCTG	1676
Qy	813	AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGCTTAAATAAGGG	872
Db	1677	AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGCTTAAATAAGGG	1736
Qy	873	GAAGAAGCGGTATGATCCCTTTGATCCAGTCACTTGAAACTGTTTACCAATCAATACGTT	932
Db	1737	GAAGAAGCGGTATGATCCCTTTGATCCAGTCACTTGAAACTGTTTACCAATCAATACGTT	1796
Qy	933	GATGTCGATCAACGAATTTGCTAAAGGTGAGCAGCTCTTAAACAGCTAGCGCAACGTAAC	992
Db	1797	GATGTCGATCAACGAATTTGCTAAAGGTGAGCAGCTCTTAAACAGCTAGCGCAACGTAAC	1856
Qy	993	TTAGACTTCAGAGATTTATACGATTCCTGATGATAAGGCTAAACTCTTACAAATCTC	1052
Db	1857	TTAGACTTCAGAGATTTATACGATTCCTGATGATAAGGCTAAACTCTTACAAATCTC	1916
Qy	1053	GATGCTTTTGGTATTATGACTATACCTTAACTGGAAGAGTAGAGATATACGATGAC	1112

Db	1917	GATGCTTTTGGTATTATGGACTATACCTTAACCTGGAAGAGTAGAGGATAATACGATGAC	1976
Qy	1113	ACCAACGGTATCATACCGTTTATATGCGCAAGCGACCCGGAAGGAGAGAAATGCTAGCTAC	1172
Db	1977	ACCAACGGTATCATACCGTTTATATGCGCAAGCGACCCGGAAGGAGAGAAATGCTAGCTAT	2036
Qy	1173	CATTAGCTGGTGGT	1187
Db	2037	CATTAGCCTATGAT	2051
RESULT 6			
SEDEXB			
LOCUS	SDSEXB	8931 bp	DNA linear BCT 18-APR-2005
DEFINITION	S.equisimilis dexb, abc, lrp, skc, rel genes and ORF1.		
ACCESSION	X72832		
VERSION	X72832.1 GI:407876		
KEYWORDS	abc gene; ABC transporter; dexB gene; dextran glucosidase; glucan 1,6-alpha-glucosidase; leucine rich protein; LRP gene; ORF1; rel gene; skc gene; streptokinase; stringent response-like protein.		
SOURCE	Streptococcus dysgalactiae subsp. equisimilis		
ORGANISM	Streptococcus dysgalactiae subsp. equisimilis		
REFERENCE	1 (bases 3621 to 6190)		
AUTHORS	Malke,H., Roe,B. and Ferretti,J.J.		
TITLE	Nucleotide sequence of the streptokinase gene from Streptococcus equisimilis H46A		
JOURNAL	Gene 34 (2-3), 357-362 (1985)		
PUBMED	2989113		
REFERENCE	2 (bases 1 to 4188; 5790 to 8931)		
AUTHORS	Mechold,U., Steiner,K., Vettermann,S. and Malke,H.		
TITLE	Genetic organization of the streptokinase region of the Streptococcus equisimilis H46A chromosome		
JOURNAL	Mol.Gen. Genet. 241 (1-2), 129-140 (1993)		
PUBMED	8232196		
REFERENCE	3 (bases 1 to 8931)		
AUTHORS	Malke,H.		
TITLE	Direct Submission		
JOURNAL	Submitted (05-MAY-1993) H. Malke, Institute for Molecular Biology, Jena University, Winzerlaer Str 10, 07708 Jena, FRG		
COMMENT	Related sequences: K02986, M19346, X13399 & X13400.		
FEATURES	Location/Qualifiers		
source	1..8931		
	/organism="Streptococcus dysgalactiae subsp. equisimilis"		
	/mol_type="genomic DNA"		
	/strain="H46A"		
	/isolate="human group C strain"		
	/sub_species="equisimilis"		
	/db_xref="taxon:119602"		
	/chromosome="streptokinase region"		
	/clone_lib="E.coli plasmid library containing subfragments of the submitted seq:pSHD14/16, pSPV19, pSH2, pMF1, pCWP73, pRH10, pWx4"		
gene	complement(89..1761)		
	/gene="dexB"		
terminator	complement(89..115)		
	/gene="dexB"		
	/note="hairpin loop"		
	complement(136..1749)		
CDS	/gene="dexB"		
	/EC_number="3.2.1.70"		
	/codon_start=1		
	/transl_table=11		
	/product="glucan 1,6-alpha-glucosidase"		
	/protein_id="CAA51348.1"		
	/db_xref="GI:407877"		
	/db_xref="GOA:Q59905"		
	/db_xref="InterPro:IPR006047"		
	/db_xref="InterPro:IPR006589"		
	/db_xref="UniProt/Swiss-Prot:Q59905"		
	/translation="MQQWVHKATIVQIPRSPKDTSGNGIGDLKGTISOLDYLQKLG ITAIWLSPTVQSPMDNGYDIDSYEALAEVFGNDDMDLLAAANERGIKIIMDLVFN		

Db	4517	ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACCAACGCCAAATAGTTGTT	4576
Qy	93	AGCGTTGCTGGTACTGTTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTTTTGGAATC	152
Db	4577	AGCGTTGCTGGTACTGTTTGAGGGGACGATCAAGACATTAAGTCTTAAATTTTTTTGGAATC	4636
Qy	153	GATCTAAATCAACGACCTGTCTATGAGGAGAAAGACAGACGAAGGCTTAAAGTCCAAATCA	212
Db	4637	GATCTAAATCAACGACCTGTCTATGAGGAGAAAGACAGACGAAGGCTTAAAGTCCAAATCA	4696
Qy	213	AAACCATTTGCTACTGTAGTAGTGGCGCGATGTCAATAAATTGAGAAAGCTGACTTACTA	272
Db	4697	AAACCATTTGCTACTGTAGTAGTGGCGCGATGTCAATAAATTGAGAAAGCTGACTTACTA	4756
Qy	273	AAGGCTATTCAAGAAACAATTGATCGCTAAACGTGCCACAGTAAACGACGACTACTTTGAGGTC	332
Db	4757	AAGGCTATTCAAGAAACAATTGATCGCTAAACGTGCCACAGTAAACGACGACTACTTTGAGGTC	4816
Qy	333	ATTGATTTTGCACGCGATGCAACATTAATGATCGAAACCGCAAGGCTCTACTTTGCTGAC	392
Db	4817	ATTGATTTTGCACGCGATGCAACATTAATGATCGAAACCGCAAGGCTCTACTTTGCTGAC	4876
Qy	393	AAAGATGGTTGGGTAAACCTTGGCCGACCCACCTGTCCAAGAAATTTTGTCTAAGCGGACAT	452
Db	4877	AAAGATGGTTGGGTAAACCTTGGCCGACCCACCTGTCCAAGAAATTTTGTCTAAGCGGACAT	4936
Qy	453	GTGGCGGTTAGACCATATAAGAAAAACCAATACAAAAACCAAGCGAAATCTGTTGATGTG	512
Db	4937	GTGGCGGTTAGACCATATAAGAAAAACCAATACAAAAACCAAGCGAAATCTGTTGATGTG	4996
Qy	513	GAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAA	572
Db	4997	GAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAA	5056
Qy	573	GATACTAAGCTATTGAAAAACATGACTTATCGGTGACACCATCATCTCAAGAAATTACTA	632
Db	5057	GATACTAAGCTATTGAAAAACATGACTTATCGGTGACACCATCATCTCAAGAAATTACTA	5116
Qy	633	GCTCAAGCACAAAGCATTTTAAACAAACCAACCCAGCGCTATACGATTTATGAACGTGAC	692
Db	5117	GCTCAAGCACAAAGCATTTTAAACAAACCAACCCAGCGCTATACGATTTATGAACGTGAC	5176
Qy	693	TCCTCAATCGTCACTCATGACAAATGACATTTTCGGTACGATTTTACCAATGGATCAAGAG	752
Db	5177	TCCTCAATCGTCACTCATGACAAATGACATTTTCGGTACGATTTTACCAATGGATCAAGAG	5236
Qy	753	TTTACTTACCGTGTAAAAATCGGGAAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG	812
Db	5237	TTTACTTACCGTGTAAAAATCGGGAAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG	5296
Qy	813	AATGAAGAAATAACACACATGACCTCATCTCTGAGAAATATTACGTCCTTAAAAAGGG	872
Db	5297	AATGAAGAAATAACACACATGACCTCATCTCTGAGAAATATTACGTCCTTAAAAAGGG	5356
Qy	873	GAAGAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAAAATGTTTACCATCAAAATACGTT	932
Db	5357	GAAGAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAAAATGTTTACCATCAAAATACGTT	5416
Qy	933	GATGTCGATACCAACGAATTTGCTTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAAC	992
Db	5417	GATGTCGATACCAACGAATTTGCTTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAAC	5476
Qy	993	TTAGACTTCAGAGNTTATACGATCCCTCGTGATTAAGGCTAAATCTCTCTACAAATCTC	1052
Db	5477	TTAGACTTCAGAGNTTATACGATCCCTCGTGATTAAGGCTAAATCTCTCTACAAATCTC	5536
Qy	1053	GATGCTTTTGGTATTATGGAATATACCTTAACTGGAAGAGTAGAGGATTAATCAGATGAC	1112
Db	5537	GATGCTTTTGGTATTATGGAATATACCTTAACTGGAAGAGTAGAGGATTAATCAGATGAC	5596
Qy	1113	ACCAACCGTATCAAAACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAAATGCTAGCTAC	1172

Db	5597	ACCAACGGTATCATACCGTTTATATGGCAAGCGACGCCAAGAGAGAATGCTAGCTAT	5655
Qy	1173	CATTAGCTGGTGCT	1187
Db	5657	CATTAGCCTATGAT	5671
RESULT 7			
LOCUS	E00522	2568 bp	DNA linear PAT 29-SEP-1997
DEFINITION	DNA fragment of plasmid PMF1 into which DNA encoding streptokinase is inserted.		
ACCESSION	E00522		
VERSION	E00522.1	GI:2168801	
KEYWORDS	JP 1985237995-A/1.		
SOURCE	Streptococcus dysgalactiae subsp. equisimilis		
ORGANISM	Streptococcus dysgalactiae subsp. equisimilis		
	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Streptococcus.		
REFERENCE	1 (bases 1 to 2568)		
AUTHORS	Jiyousefu, J F, and Horusuto, M.		
TITLE	RECOMBINATION VECTOR FOR PRODUCING STREPTOKINASE		
JOURNAL	Patent: JP 1985237995-A 1 26-NOV-1985;		
COMMENT	PHILIPS PETROLEUM CO		
	OS Streptococcus equisimilis		
	PN JP 1985237995-A/1		
	PD 26-NOV-1985		
	PF 09-OCT-1984 JP 1984212403		
	PR 10-OCT-1983 DD 83 255523, 02-MAR-1984 US 84 585417 PI		
	JIYOSUFU JIBI FUERTSUTETEI, HORUSUTO MARUKE		
	PC C12N15/00, C12N1/20, C12N9/70, (C12N15/00, C12R1:46), (C12N1/20, PC C12R1:19),		
	PC (C12N9/70, C12R1:19);		
	CC strandedness: Single;		
	CC topology: Linear;		
	CC hypothetical: No;		
	CC anti-sense: No;		
	CC *source: strain=Streptococcus equisimilis H46A; CC *source: library=streptococcus equisimilis H46 library; CC *source: clone=lambda L47 skc clone;		
	FH Key	Location/Qualifiers	
	FT 5'UTR	1..818	
	FT sig_peptide	819..896	
	FT	/product='streptokinase signal peptide' FT	
	mat_peptide	897..2138	
	FT	/product='streptokinase'	
	FT CDS	819..2143	
	FT	/product='streptokinase'	
	FT 3'UTR	2142..2569.	
FEATURES		Location/Qualifiers	
source	1..2568		
	/organism="Streptococcus dysgalactiae subsp. equisimilis"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:119602"		
ORIGIN			
Query Match	73.2%	Score 1127.8; DB 6; Length 2568;	
Best Local Similarity	98.5%	Pred. No. 5.2e-277;	
Matches 1138;	Conservative 0; Mismatches 17; Indels 0; Gaps 0;		
Qy	33	ATAGCTGGTCTCGAATGGCTACTAGATCGTCCCTTCTGTAATAACAGCCAAATGGTTGTT	92
Db	897	ATTGCTGGACTGATGGCTGCTAGACCGTCCATCTGTCAACACGCAATAGTTGTT	956
Qy	93	AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTTAAATTTTGAATC	152
Db	957	AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTTAAATTTTGAATC	1016
Qy	153	GATCTAACATCAAGCACTGCTCATGGAGGAAGACAGACAGCGCTTAAGTCCAAATCA	212
Db	1017	GATCTAACATCAAGCACTGCTCATGGAGGAAGACAGACAGCGCTTAAGTCCAAATCA	1076

Qy	213	AAACCATTTTGGTACTGATAGTGGCGCGCATGTGCACATAAACTTTGAGAAAGCTGACTTACTTA	272
Db	1077		
Qy	273	AAAGCTATTCAAGAAACAATTTGATCGCTAAACGTCCACAGTAGTAACGACGACTACTTTTGAGGTC	332
Db	1137		
Qy	333	ATTGATTTTGCAGCGATGCAACCAATTTACTGATCGAAACCGCAAGGCTCTACTTTGCTGCAC	392
Db	1197		
Qy	393	AAAGATGGTTCCGGTAAACCTTTCGCGACCCCAACCTGTCCAAAGATTTTTCGTAAAGCGGACAT	452
Db	1257		
Qy	453	GTGCGCGTTAGACCATATAAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG	512
Db	1317		
Qy	513	GAATATACTGTACAGTTTATCTCCCTTTAAACCCCTGATGACGATTTTCAGACCAAGGTCCTCAAA	572
Db	1377	GAATATATACTGTACAGTTTATCTCCCTTTAAACCCCTGATGACGATTTTCAGACCAAGGTCCTCAAA	632
Qy	573	GATTAAGCTATTGAAAAACATAGCTTATCGGTGACACCAATCAATCTCAAGAAATTACTTA	632
Db	1437	GATTAAGCTATTGAAAAACATAGCTTATCGGTGACACCAATCAATCTCAAGAAATTACTTA	1496
Qy	633	GCTCAAGCACAAGCATTTTAAACAAAAACCAACCCAGGCTATACGATTTATGAACGTGAC	692
Db	1497	GCTCAAGCACAAGCATTTTAAACAAAAACCAACCCAGGCTATACGATTTATGAACGTGAC	1556
Qy	693	TCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTAAACCAATGCGATCAAGAG	752
Db	1557	TCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTAAACCAATGCGATCAAGAG	1616
Qy	753	TTTACTTACCCTGTTTAAAAATCGSGAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG	812
Db	1617	TTTACTTACCCTGTTTAAAAATCGSGAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG	1676
Qy	813	AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAACGTCCTTAAAAAAGGG	872
Db	1677	AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAACGTCCTTAAAAAAGGG	1736
Qy	873	GAAGAAGCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTTCACCATCAAAATACGTT	932
Db	1737	GAAGAAGCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTTCACCATCAAAATACGTT	1796
Qy	933	GATGTCGATACCAACGAATTTGCTTAAAGATGAGCAGCTCTTTAAACAGCTAGCGAACGTTAAC	992
Db	1797	GATGTCGATACCAACGAATTTGCTTAAAGATGAGCAGCTCTTTAAACAGCTAGCGAACGTTAAC	1856
Qy	993	TTAGACTTCAGAGATTTTATACGATCTCTCGTGTAAGGCTAAACTCTCTCAACAATCTC	1052
Db	1857	TTAGACTTCAGAGATTTTATACGATCTCTCGTGTAAGGCTAAACTCTCTCAACAATCTC	1916
Qy	1053	GATGCTTTTCGGTATTATGGACTATACCTTAACTGGAAAAAGTAGGAGTATACGATGAC	1112
Db	1917	GATGCTTTTCGGTATTATGGACTATACCTTAACTGGAAAAAGTAGGAGTATACGATGAC	1976
Qy	1113	ACCAACCGTATCATAAACCGTTTATATGGCGAAAGCGACCCGAAGGAGAGATGCTAGCTAC	1172
Db	1977	ACCAACCGTATCATAAACCGTTTATATGGCGAAAGCGACCCGAAGGAGAGATGCTAGCTAC	2036
Qy	1173	CATTTAGCTGGTGGT	1187
Db	2037	CATTTAGCTTATGAT	2051

RESULT 8			
C0797820			
LOCUS	C0797820	7057 bp	DNA
DEFINITION	Sequence 26 from Patent WO2004029256.		
		linear	PAT 20-APR-2004

ACCESSION	CQ797820
VERSION	CQ797820.1 GI:46426093
KEYWORDS	.
SOURCE	synthetic construct
ORGANISM	other sequences; artificial sequences.
REFERENCE	1
AUTHORS	selman-housain Sosa,G., aguiar Cabeza,B., gonz lez Quintero,A.D. y ramos gonz Lez,O.
TITLE	Vector for the production of transplastomic angiosperm plants
JOURNAL	Patent: WO 2004029256-A 26 08-APR-2004;
FEATURES	CENTRO DE INGENIERIA GENETICA Y BIOTECNOLOGIA (CU)
source	Location/Qualifiers 1..7057 /organism="synthetic construct" /mol_type="unassigned DNA" /db_xref="taxon:32630" /note="Description of Artificial Sequence: Nucleotide sequence of DNA fragment from the vector pVTPA-Estrep between the rice atpB and tobacco rbcL borders."
ORIGIN	
Query Match	72.7%; Score 1120.2; DB 6; Length 7057;
Best Local Similarity	97.6%; Pred.No. 4.5e-275;
Matches 1137; Conservative	0; Mismatches 28; Indels 0; Gaps 0;
Qy	23 ATATACCATTGATAGCTGGTCTCGTAATGGCTACTAGATCGTCCTTGTTGAATAAACAGCCA 82
Db	2013 ATAACAATGATTGTCTGGAACTGATGGCTGTCTAGACCGTCCATCTGTCTAACCAACAGCCA 2072
Qy	83 ATTGGTTGTTAGCGTTGCTCGTACTGTGTGAGGGGACGAATCAAGACATTAGTCTTAAATT 142
Db	2073 ATTAGTTGTTAGCGTTGCTGTACTGTGTGAGGGGACGAATCAAGACATTAGTCTTAAATT 2132
Qy	143 TTTTGAATAATCGATCTAAATCAGCATCAGACTGCTGCATGTGAGGAGAAGACAGACGACGCTTAAG 202
Db	2133 TTTTGAATAATTGACCTAAATCATCACGACGCTGCTCATGTGAGGAGAAGACAGACGACGCTTAAG 2192
Qy	203 TCACAAATCAAACCAATTTCTACTCATGTGCGGGATGTCTACATAAACTTCAGAAAAGC 262
Db	2193 TCACAAATCAAACCAATTTCTACTCATGTGCGGGATGCCACATAAACTTCGAAAAGC 2252
Qy	263 TGACTTACTAAAGGCTATTCAAGAACAAATTGATCGCTTAACGCTCCACAGTAACGACGACTA 322
Db	2253 TGACTTACTAAAGGCTATTCAAGAACAAATTGATCGCTTAACGCTCCACAGTAACGACGACTA 2312
Qy	323 CTTTGAGGTCATTGATTTTGCAGACGGANTGCAACATTACTGATCGAAAAACGGCAAGGTCCTA 392
Db	2313 CTTTGAGGTCATTGATTTTGCAGACGGATGCAACCAATTACTGATCGAAAAACGGCAAGGTCCTA 2372
Qy	383 CTTTGTCTGACAAAGATGGTTTCGGTAACTTGCCTGCGACCCCAACCTGTCACCAAGATTTTGGCT 442
Db	2373 CTTTGTCTGACAAAGATGGTTTCGGTAACTTGCCTGCGACCCCAACCTGTCACCAAGATTTTGGCT 2432
Qy	443 AAGCGGACATGTGCGCGCTTAGACCATATAAAGAAAAACCAATACAAAAACCAACGCGAAATC 502
Db	2433 AAGCGGACATGTGCGCGCTTAGACCATATAAAGAAAAACCAATACAAAAACCAACGCGAAATC 2492
Qy	503 TGTGTGATGTGGAATATATCTGTA CAGTTTTACTTCCTTAAACCCCTGATGACGATTCAGACC 562
Db	2493 TGTGTGATGTGGAATATATCTGTA CAGTTTTACTTCCTTAAACCCCTGATGACGATTCAGACC 2552
Qy	563 AGGTCTCAAGATACTTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCATCATCTCA 622
Db	2553 AGGTCTCAAGATACTTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCATCATCTCA 2612
Qy	623 AGAATTACTAGCTCAAGCACAAAGCAATTTTAAACAAAAACCAACCCAGGCTATACGATTTA 682
Db	2613 AGAATTACTAGCTCAAGCACAAAGCAATTTTAAACAAAAACCAACCCAGGCTATACGATTTA 2672
Qy	683 TGAAACGTGACTCCTCAATCGTCACTCATGACAATGTACATTTCCCGTACGATTTTACCAAT 742
Db	2673 TGAAACGTGACTCCTCAATCGTCACTCATGACAATGTACATTTCCCGTACGATTTTACCAAT 2732

KEYWORDS

SOURCE Streptococcus dysgalactiae subsp. equisimilis
 ORGANISM Streptococcus dysgalactiae subsp. equisimilis
 Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 Streptococcus.

REFERENCE

AUTHORS Estrada,M.P., Hernandez,L., Perez,A., Rodriguez,P., Serrano,R.,
 Rubiera,R., Pedraza,A., Padron,G., Antuch,W., de la Fuente, J. et
 al.

TITLE High level expression of streptokinase in Escherichia coli

JOURNAL Biotechnology (N.Y.) 10 (10), 1138-1142 (1992)

PUBMED 1368792

REMARK GenBank staff at the National Library of Medicine created this
 entry [NCBI gibbsg 115306] from the original journal article.

FEATURES

source
 1..1245
 /organism="Streptococcus dysgalactiae subsp. equisimilis"
 /mol_type="genomic DNA"
 /db_xref="taxon:119602"
 /note="group: C"
 gene
 1..1245
 /gene="SKC-2"
 CDS
 1..1245
 /gene="SKC-2"
 /note="streptokinase; Mature protein lacks initial Met"
 /codon_start=1
 /product="SKC-2"
 /protein_id="AAC60418.1"
 /db_xref="GI:257197"
 /translation="TAGPEWLLDRPSVNNLSQVSVAGTVVEGTNODISLKFPFIDILTS
 RPAHGKTEQGLSPKSPATDSGAMPHKLEKADLLKATQEOILIANVHNDYDVFVID
 PASDAITIDRNGKVVFADKQGSVLTPTQVQFLLSGHVRVPYKEPKIQNOAKSVDV
 EYTVQFTPLNPDPPRGLKDTLLKTLAIGDTITSOELLAQAQSLNKHPTGYTIE
 LKSSVITHNDIFRILPMQBFTHVKNREQAYEINKKSGLEENINNTDLISEKYYV
 RDGKPKYDPPDRSLKLFITKYVDVNTNELLSQLLTASERNLDFRDLYPDPRDCAK
 LLYNLDLAFGIMDYTLTKVEDNHDDTNRIITVTKRPEGENASVHLAYDKRYTEE
 EREVSYLRYTGTPIPDNDK"

ORIGIN

Query Match 72.1%; Score 1111.8; DB 1; Length 1245;

Best Local Similarity 97.7%; Pred. No. 6.5e-273;

Matches 1128; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 33 ATAGCTGTCCTGAATGGCTACTAGATCGTCTTCGTAAATAACAGCCAAATGGTGT 92
 DB 1 ATTGCTGCACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCCAAATGGTGT 60
 QY 93 AGCGTTGCTGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGGAAATC 152
 DB 61 AGCGTTGCTGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGGAAAT 120
 QY 153 GATCTAAATCAACGACCTGCTCATGGAGGAAAGACAGACAAAGCTTAAAGTCCAAATCA 212
 DB 121 GACCTAAATCAACGACCTGCTCATGGAGGAAAGACAGACAAAGCTTAAAGTCCAAATCA 180
 QY 213 AAACCAATTTGCTACTGATAGTGGCGGATGTACATAAACTTGAGAAAGCTACTTACTA 272
 DB 181 AAACCAATTTGCTACTGATAGTGGCGGATGCCACATAAACTTGAGAAAGCTACTTACTA 240
 QY 273 AAGGCTATTCAAGAACAAATTTGATCGTAACTGCTCAAGTAAACGACGACTACTTTGAGTTC 332
 DB 241 AAGGCTATTCAAGAACAAATTTGATCGTAACTGCTCAAGTAAACGACGACTACTTTGAGTTC 300
 QY 333 ATTGATTTTGAAGCGATGCAACCAATCTACTGATCGAAACGGCAAGGTCTACTTTGCTGAC 392
 DB 301 ATTGATTTTGAAGCGATGCAACCAATCTACTGATCGAAACGGCAAGGTCTACTTTGCTGAC 360
 QY 393 AAAGATGTTTGGTAACCTTGGCGACCCAACTGTCCTCAAGAAATTTTGTCTAAGCGACAT 452
 DB 361 AAAGATGTTTGGTAACCTTGGCGACCCAACTGTCCTCAAGAAATTTTGTCTAAGCGACAT 420
 QY 453 GTGCGCGTTAGACCATATATAAGAAACCAATATACAAACCAAGCGAAATCTGTTGATGTG 512

DB 421 GTGCGGTTAGACCATTATAAGAAAAACCAATACAAATCAAGCGAAATCTGTTGATGTG 480
 QY 513 GAATATACCTGACAGTTTACTCCCTTAAACCCCTGATGACGATTTCACACAGAGTCTCAAA 572
 DB 481 GAATATACCTGACAGTTTACTCCCTTAAACCCCTGATGACGATTTCACACAGAGTCTCAAA 540
 QY 573 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCCATCATCATCTCAAGAATTACTA 632
 DB 541 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCCATCATCATCTCAAGAATTACTA 600
 QY 633 GCTCAAGCAACAAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTATGAACGTCGAC 692
 DB 601 GCTCAAGCAACAAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTATGAACGTCGAC 660
 QY 693 TCCTCAATCGTCACTCATGACAAATGACATTTTCGTCGATCAATTTTACCAATGGATCAAGAG 752
 DB 661 TCCTCAATCGTCACTCATGACAAATGACATTTTCGTCGATCAATTTTACCAATGGATCAAGAG 720
 QY 753 TTTTACTTACCGTGTAAATAATCGGAAACAAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 812
 DB 721 TTTTACTTACCGTGTAAATAATCGGAAACAAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 780
 QY 813 AATGAAGAAATAAACCAACACTGACCTGATCTCTGAGAAATATTAACGTCCTTAAAAAAGG 872
 DB 781 AATGAAGAAATAAACCAACACTGACCTGATCTCTGAGAAATATTAACGTCCTTAAAAAAGG 840
 QY 873 GAAAGCGGTATGATCCCTTTGATCGAGTCACTTGAAACTGTTCCACCATCAATAGCTT 932
 DB 841 GAAAGCGGTATGATCCCTTTGATCGAGTCACTTGAAACTGTTCCACCATCAATAGCTT 900
 QY 933 GATGTCGATACCAACCAATTTGCTAAAAAGTGACGAGCTCTTAAACAGCTAGCGAACGTAAC 992
 DB 901 GATGTCGATACCAACCAATTTGCTAAAAAGTGACGAGCTCTTAAACAGCTAGCGAACGTAAC 960
 QY 993 TTAGACTTCAGAGATTTATACGATCTCTGATGATGAGCTAAAGCTAAACCTCTCAACAATCTC 1052
 DB 961 TTAGACTTCAGAGATTTATACGATCTCTGATGATGAGCTAAAGCTAAACCTCTCAACAATCTC 1020
 QY 1053 GATGCTTTTGGTATTATGAGCTATACCTTAACTGGAAAAAGTAGAGGATATACGATGAC 1112
 DB 1021 GATGCTTTTGGTATTATGAGCTATACCTTAACTGGAAAAAGTAGAGGATATACGATGAC 1080
 QY 1113 ACCAACCGTATCATACCGTTTATATGCGCAACGCGCTAAAGCTAAAGCTAAAGCTAGCTAC 1172
 DB 1081 ACCAACCGTATCATACCGTTTATATGCGCAACGCGCTAAAGCTAAAGCTAGCTAGCTAT 1140
 QY 1173 CATTTAGCTGGTGGT 1187
 DB 1141 CATTTAGCTGATGAT 1155

RESULT 11

A20015
 LOCUS SEQ ID NO: 23; Nucleotide sequence for methionol-streptokinase
 DEFINITION fusion protein.
 ACCESSION A20015
 VERSION A20015.1 GI:1247848
 KEYWORDS synthetic construct
 SOURCE synthetic construct
 ORGANISM other sequences; artificial sequences.
 REFERENCE 1 (bases 1 to 1257)
 AUTHORS PROTEINS AND NUCLEIC ACIDS
 TITLE Patent: WO 9109125-A 23 27-JUN-1991;
 JOURNAL Location/Qualifiers
 FEATURES
 source
 1..1257
 /organism="synthetic construct"
 /mol_type="unassigned DNA"
 /db_xref="taxon:32630"
 4..1251
 /codon_start=1
 CDS

/transl_table=11
/product="methionyl-streptokinase fusion protein"
/protein_id="CAA01486.1"
/db_xref="GI:1247849"
/translation="MTAGPEWLLDRPSVNNLSQVSVAGTVENQDLSIKPEIDL
SPAHGKTEQGLSPKSPKPDATSGAMPKLEKADLLKALQELIANVHNDYFEVI
DFASDTITDRNGKVFPADKDGSVTLTPQBPFLISGHVVRPYKEKPIQNAQSV
VRYTQFTPLNPDPPRGLKDTLQAGDTTSQELLAQSLINXTHFGYTIY
ERDSSTVINDIIFRILPMDQFTVYKVRNEQAYEINKKSLNEEINNDLISEKY
VLKSGKPYDPDRSHLFTIKYVDVNTNELKSEOLLTASERNLDFRDLXPRDKA
KLLYNLDAFGIMDYTLTGKVEDNDDTNRIIVYMKRPEGENASHLAYDKDRTE
EREVYSYLRYGTPTPDNDK"

ORIGIN

Query Match 72.1%; Score 1111.6; DB 6; Length 1257;
Best Local Similarity 97.5%; Pred. No. 7.3e-273;
Matches 1129; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 30 ATGATAGCTGGTCTCAATGGCTACTAGATCGCTCTCTTAATAACAGCCCAATTGGTT 89
DB 4 ATGATAGCTGGTCTCAATGGCTACTAGATCGCTCTCTTAATAACAGCCCAATTGGTT 63

QY 90 GTTAGCGTTGCTGGTACTGTGAGGGGACGAATCAAGACATTAAGTCTTAATTTTGGAA 149
DB 64 GTTAGCGTTGCTGGTACTGTGAGGGGACGAATCAAGACATTAAGTCTTAATTTTGGAA 123

QY 150 ATCGATCTAACATCACGACCTGCTCATGGAGGAAAGACAGACGAGCTTTAAGTCCAAAA 209
DB 124 ATTAGCTTAACATCACGACCTGCTCATGGAGGAAAGACAGACGAGCTTTAAGTCCAAAA 183

QY 210 TCAAAACCAATTTGCTACTGATAGTGGCGGATGTCATAAATCGAGAAAGCTGACTTA 269
DB 184 TCAAAACCAATTTGCTACTGATAGTGGCGGATGTCATAAATCGAGAAAGCTGACTTA 243

QY 270 CTAAGGCTATTCAAGACAAATTTGATCGCTAAACGTCACAGTAAACGACGACTTTGAG 329
DB 244 CTAAGGCTATTCAAGACAAATTTGATCGCTAAACGTCACAGTAAACGACGACTTTGAG 303

QY 330 GTCAATGATTTTGCAGCGATGCAACCATTTACTGATCGAAACGCAAGTCTACTTTGCT 389
DB 304 GTCAATGATTTTGCAGCGATGCAACCATTTACTGATCGAAACGCAAGTCTACTTTGCT 363

QY 390 GACAAAGATGGTTGGTAACTTGGCGGACCAACCTGTCCAGAAATTTTGTAAAGCGGA 449
DB 364 GACAAAGATGGTTGGTAACTTGGCGGACCAACCTGTCCAGAAATTTTGTAAAGCGGA 423

QY 450 CATGTGCGGTAGACCATATAAGAAACCAATACAAACCAAGCGAAATCTGTTGAT 509
DB 424 CATGTGCGGTAGACCATATAAGAAACCAATACAAACCAAGCGAAATCTGTTGAT 483

QY 510 GTGGAATATCTGTACAGTTTACTCCCTTAACCCCTGATGACGATTCAGACAGGCTCTC 569
DB 484 GTGGAATATCTGTACAGTTTACTCCCTTAACCCCTGATGACGATTCAGACAGGCTCTC 543

QY 570 AAGATATCTAAGCTATTGAAACACATGCTATCGGTGACACCATCATCTCAAGAAATTA 629
DB 544 AAGATATCTAAGCTATTGAAACACATGCTATCGGTGACACCATCATCTCAAGAAATTA 603

QY 630 CTAGCTCAAGCAACAGCAATTTTAAACAAAAACCCAGGCTATACGATTTATGAACGT 689
DB 604 CTAGCTCAAGCAACAGCAATTTTAAACAAAAACCCAGGCTATACGATTTATGAACGT 663

QY 690 GACTCTCAATCGTCTCATGACAAATGACATTTTCCGTCAGATTTTACCAATGGATCAA 749
DB 664 GACTCTCAATCGTCTCATGACAAATGACATTTTCCGTCAGATTTTACCAATGGATCAA 723

QY 750 GAGTTTACTTACCGTTTAAATCGGAAACAGCTTATAGGATCAATAAAAAATCTGGT 809
DB 724 GAGTTTACTTACCGTTTAAATCGGAAACAGCTTATAGGATCAATAAAAAATCTGGT 783

QY 810 CTGAATGAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTAACGCTCTTAAAAA 869
DB 784 CTGAATGAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTAACGCTCTTAAAAA 843

QY 870 GGGGAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCCACCATCAAAATAC 929
DB 844 GGGGAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCCACCATCAAAATAC 903

QY 930 GTTGATGTGATACCAACGAATTCCTAAAAGTGAAGAGTCTTTAAACAGCTAGCGAAAGT 989
DB 904 GTTGATGTCAACACCAACGAATTCCTAAAAGTGAAGAGTCTTTAAACAGCTAGCGAAAGT 963

QY 990 AACCTAGACTTTCAGAGATTTATACGATCCCTGCTGATAGGCTAACTCTTACACAAAT 1049
DB 964 AACCTAGACTTTCAGAGATTTATACGATCCCTGCTGATAGGCTAACTCTTACACAAAT 1023

QY 1050 CTCGATGCTTTTGGTATTATGGAATATACCTTTAACTGMAAAGTAGAGGATAATCACCAT 1109
DB 1024 CTCGATGCTTTTGGTATTATGGAATATACCTTTAACTGMAAAGTAGAGGATAATCACCAT 1083

QY 1110 GACACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAGAGAGAAATGCTAGC 1169
DB 1084 GACACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAGAGAGAAATGCTAGC 1143

QY 1170 TACCATTAGCTGGTGGT 1187
DB 1144 TATCATTTAGCCTATGAT 1161

RESULT 12
I13203
LOCUS I13203 1257 bp DNA linear PAT 26-JUL-1995
DEFINITION Sequence 25 from patent US 5434073.
ACCESSION I13203
VERSION I13203.1 GI:910551
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1257)
AUTHORS Dawson,K., Hunter,M.G. and Czaplowski,L.G.
TITLE Fibrinolytic and anti-thrombotic cleavable dimers
JOURNAL Patent: US 5434073-A 25 JUL-1995;
FEATURES Location/Qualifiers
source 1..1257
/organism="unknown"
/mol_type="unassigned DNA"

Query Match 72.1%; Score 1111.6; DB 6; Length 1257;
Best Local Similarity 97.5%; Pred. No. 7.3e-273;
Matches 1129; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 30 ATGATAGCTGGTCTCAATGGCTACTAGATCGCTCTCTTAATAACAGCCCAATTGGTT 89
DB 4 ATGATAGCTGGTCTCAATGGCTACTAGATCGCTCTCTTAATAACAGCCCAATTGGTT 63

QY 90 GTTAGCGTTGCTGGTACTGTGAGGGGACGAATCAAGACATTAAGTCTTAATTTTGGAA 149
DB 64 GTTAGCGTTGCTGGTACTGTGAGGGGACGAATCAAGACATTAAGTCTTAATTTTGGAA 123

QY 150 ATCGATCTAACATCACGACCTGCTCATGGAGGAAAGACAGACGAGCTTTAAGTCCAAAA 209
DB 124 ATTAGCTTAACATCACGACCTGCTCATGGAGGAAAGACAGACGAGCTTTAAGTCCAAAA 183

QY 210 TCAAAACCAATTTGCTACTGATAGTGGCGGATGTCATAAATCGAGAAAGCTGACTTA 269
DB 184 TCAAAACCAATTTGCTACTGATAGTGGCGGATGTCATAAATCGAGAAAGCTGACTTA 243

QY 270 CTAAGGCTATTCAAGACAAATTTGATCGCTAAACGTCACAGTAAACGACGACTTTGAG 329
DB 244 CTAAGGCTATTCAAGACAAATTTGATCGCTAAACGTCACAGTAAACGACGACTTTGAG 303

QY 330 GTCAATGATTTTGCAGCGATGCAACCATTTACTGATCGAAACGCAAGTCTACTTTGCT 389
DB 304 GTCAATGATTTTGCAGCGATGCAACCATTTACTGATCGAAACGCAAGTCTACTTTGCT 363

QY	390	GACAAAGATGGTTCGGTAACTTGGCCGACCCACCTGTCCAAAGAAATTTTGTCTAAGCGGA	449
Db	364	GACAAAGATGGTTCGGTAACTTGGCCGACCCACCTGTCCAAAGAAATTTTGTCTAAGCGGA	423
QY	450	CATGTGCGCGTTAGACCATATAAAGAAAAACCAATAACAAACCAAGCGAAATCTGTGTGAT	509
Db	424	CATGTGCGCGTTAGACCATATAAAGAAAAACCAATAACAAACCAAGCGAAATCTGTGTGAT	483
QY	510	GTGGGAATATACCTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCGGTCTC	569
Db	484	GTGGGAATATACCTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCGGTCTC	543
QY	570	AAAGATATCTAAGCTATTGAAACACATAGCTATCGGTGACACCATCAATCTCAAGAAATTA	629
Db	544	AAAGATATCTAAGCTATTGAAACACATAGCTATCGGTGACACCATCAATCTCAAGAAATTA	603
QY	630	CTAGCTCAAGCAAAAGCAATTTTAAACAAACCCAGCGCTATACGATTTTATGAACGT	689
Db	604	CTAGCTCAAGCAAAAGCAATTTTAAACAAACCCAGCGCTATACGATTTTATGAACGT	663
QY	690	GACTCTCAATCGTCACTCATGACATGACATTTTCGGTACGATTTTACCAATGGATCAA	749
Db	664	GACTCTCAATCGTCACTCATGACATGACATTTTCGGTACGATTTTACCAATGGATCAA	723
QY	750	GAGTTTACTTACCGTGTAAACATCGGGAACAGCTTATAGGATCAATAAAAAATCTGGT	809
Db	724	GAGTTTACTTACCGTGTAAACATCGGGAACAGCTTATAGGATCAATAAAAAATCTGGT	783
QY	810	CTGAATGAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTACCGTCTTAAAAAA	869
Db	784	CTGAATGAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTACCGTCTTAAAAAA	843
QY	870	GGGAAAAAGCCGTATGATCCCTTTGATCGGAGTCACTTGAAACTGTTCACCAATCAAATAC	929
Db	844	GGGAAAAAGCCGTATGATCCCTTTGATCGGAGTCACTTGAAACTGTTCACCAATCAAATAC	903
QY	930	GTTGATGTCGATACCAACAAATTTGCTTAAAGTGAGCAGCTCTTAAACGTCAGCAACCT	989
Db	904	GTTGATGTCGATACCAACAAATTTGCTTAAAGTGAGCAGCTCTTAAACGTCAGCAACCT	963
QY	990	AACCTAGACTTCAGAGATTTATACGATCCTCGTGATAGGCTAAACTACTCTACAACAAT	1049
Db	964	AACCTAGACTTCAGAGATTTATACGATCCTCGTGATAGGCTAAACTACTCTACAACAAT	1023
QY	1050	CTCGATGCTTTTGGTATTATGACATATACCTTAACCTGGAAGAGTAGAGGATATCAGAT	1109
Db	1024	CTCGATGCTTTTGGTATTATGACATATACCTTAACCTGGAAGAGTAGAGGATATCAGAT	1083
QY	1110	GACACCAACCGTATCATACCGTTTATATGGGCAAGCAGCCGAGGAGAGATGCTAGC	1169
Db	1084	GACACCAACCGTATCATACCGTTTATATGGGCAAGCAGCCGAGGAGAGATGCTAGC	1143
QY	1170	TACCATTTAGCTGGTGGT 1187	
Db	1144	TATCATTTAGCTGATGAT 1161	
RESULT 13			
LOCUS	E01413	1407 bp	DNA linear PAT 29-SBP-1997
DEFINITION	DNA sequence of streptokinase.		
ACCESSION	E01413		
VERSION	E01413.1	GI:2169669	
KEYWORDS	JP 1987296881-A/1.		
SOURCE	Lycopersicon esculentum (Solanum lycopersicum)		
ORGANISM	Lycopersicon esculentum		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.		
REFERENCE	1 (bases 1 to 1407)		
AUTHORS	Marii,J.H. and Deebuitsudo U.S.		
TITLE	YEAST DNA FRAGMENT AND ITS UTILIZATION		
JOURNAL	Patent: JP 1987296881-A 1 24-DEC-1987;		
COMMENT	PHILLIPS PETROLEUM CO		
PN	JP 1987296881-A/1		
PD	24-DEC-1987		
PF	02-MAY-1987 JP 1987109620		
PR	08-MAY-1986 US 86 860960		
PI	MARII JIEIN HEIJENSON, DEEBUITSUO UOOMATSUKU SUTOROOWAN PC		
CI2N15/00, CI2N1/16, CI2N9/12, (CI2N1/16, CI2R1/84); CC	strandedness:		
Double;			
CC	topology: Linear;		
CC	hypothetical: No;		
CC	anti-sense: No;		
FH	Key		
FH	Location/Qualifiers		
FT	CDS		
FT	8..1255		
FT	/product='streptokinase'		
FT	5'UTR		
FT	1..7		
FT	3'UTR		
FT	1256..1407.		
FEATURES	Location/Qualifiers		
source	1..1407		
	/organism='Lycopersicon esculentum'		
	/mol_type='genomic DNA'		
	/db_xref='taxon:4081'		
ORIGIN			
	Query Match	72.0%;	Score 1110.2; DB 6; Length 1407;
	Best Local Similarity	97.8%;	Pred. No. 1.7e-272;
	Matches 1130; Conservative	0; Mismatches 33; Indels	0; Gaps 0;
QY	25	ATACCATGATGAGCTGGTCTCTGAATGGCTACTAGATCGTCTCTCTGTAAAATAACAGCCCAAT	84
Db	3	ATTCCATGATGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCCCAAT	62
QY	85	TGTTGTTAGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGAATTTAGTCTTTAAATTTT	144
Db	63	TAGTTGTTAGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGAATTTAGTCTTTAAATTTT	122
QY	145	TTGAAATCGATCTAAACATCAGACCTGCTCATCGAGGAAGACAGAGCAAGGCTTAAGTC	204
Db	123	TTGAAATCGATCTAAACATCAGACCTGCTCATCGAGGAAGACAGAGCAAGGCTTAAGTC	182
QY	205	CAAAATCAAAACCAATTTGCTACTGATAGTGGCGCGATGTCAATAAACTTGAGAAAGCTG	264
Db	183	CAAAATCAAAACCAATTTGCTACTGATAGTGGCGCGATGTCAATAAACTTGAGAAAGCTG	242
QY	265	ACTTACTAAAGGCTATTCAAGAAACAATTGATCGCTAACGTCACAGTAAACGAGCTACT	324
Db	243	ACTTACTAAAGGCTATTCAAGAAACAATTGATCGCTAACGTCACAGTAAACGAGCTACT	302
QY	325	TTGAGGTCATTTGATTTTTCGAGCGATGCAACCATTTACTGATCGAAACGGCAAGGCTTACT	384
Db	303	TTGAGGTCATTTGATTTTTCGAGCGATGCAACCATTTACTGATCGAAACGGCAAGGCTTACT	362
QY	385	TTGCTGCAAAAGATGTTTCGTTAACTTTCGCGACCCCAACCTGTCCAAAGAAATTTTGTCTAA	444
Db	363	TTGCTGCAAAAGATGTTTCGTTAACTTTCGCGACCCCAACCTGTCCAAAGAAATTTTGTCTAA	422
QY	445	GCGGACATGTGCGGTTAGACCATATAAAGAAAAACCAATAACAAACCAAGCGAAATCTG	504
Db	423	GCGGACATGTGCGGTTAGACCATATAAAGAAAAACCAATAACAAACCAAGCGAAATCTG	482
QY	505	TTGATGTGGAATATACGTACGTTTACTCCCTTAAACCCCTGATGAGGATTTTCAGACCAG	564
Db	483	TTGATGTGGAATATACGTACGTTTACTCCCTTAAACCCCTGATGAGGATTTTCAGACCAG	542
QY	565	GTCTCAAGATACCTAAGCTATTTCGAAACACTAGCTATCGGTGACACCATCACATCTCAAG	624
Db	543	GTCTCAAGATACCTAAGCTATTTCGAAACACTAGCTATCGGTGACACCATCACATCTCAAG	602
QY	625	AATTACTAGCTCAAGCACAAGCAATTTTAAACAAAAACCAACCCAGGCTATACGATTTATG	684
Db	603	AATTACTAGCTCAAGCACAAGCAATTTTAAACAAAAACCAACCCAGGCTATACGATTTATG	662


```
QY 685 AACGTGACTCTCAATCGTCACTCATGACAATGACATATTTCCGTACGATTTTACCAATGG 744
|
|
|
Db 663 AACGTGACTCTCAATCGTCACTCATGACAATGACATATTTCCGTACGATTTTACCAATGG 722
|
|
|
QY 745 ATCAAGAGTTTACTTACCGTGTAAAAATCGGGAACAAGCTTATAGGATCAATAAAAAAT 804
|
|
|
Db 723 ATCAAGAGTTTACTTACCGTGTAAAAATCGGGAACAAGCTTATAGGATCAATAAAAAAT 782
|
|
|
QY 805 CTGGTCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTTACGTCCTTA 864
|
|
|
Db 783 CTGGTCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTTACGTCCTTA 842
|
|
|
QY 865 AAAAGGGGAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCAACCATCA 924
|
|
|
Db 843 AAAAGGGGAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCAACCATCA 902
|
|
|
QY 925 AATAGTGTGATGTCGATACCAAGAAATGCTTAAAAAGTGAAGAGCTTTAAACAGCTAGCG 984
|
|
|
Db 903 AATAGTGTGATGTCGATACCAAGAAATGCTTAAAAAGTGAAGAGCTTTAAACAGCTAGCG 962
|
|
|
QY 985 AAGCTAACTTACACTTCAGAGATTTATACGATCTCTGATAGGCTTAACTACTCTACA 1044
|
|
|
Db 963 AAGCTAACTTACACTTCAGAGATTTATACGATCTCTGATAGGCTTAACTACTCTACA 1022
|
|
|
QY 1045 ACAATCTCGATCTTTTGGTATTTATGAGCTATATACCTTAACTGGAAGTAGAGATAATC 1104
|
|
|
Db 1023 ACAATCTCGATCTTTTGGTATTTATGAGCTATATACCTTAACTGGAAGTAGAGATAATC 1082
|
|
|
QY 1105 AGATGACACCAACCGTATCATAAACGTTTATATGGCAAGCGACCGAAGGAGAGAAATG 1164
|
|
|
Db 1083 AGATGACACCAACCGTATCATAAACGTTTATATGGCAAGCGACCGAAGGAGAGAAATG 1142
|
|
|
QY 1165 CTAGCTACATTTAGCTGGTGGT 1187
|
|
|
Db 1143 CTAGCTACATTTAGCTGGTGGT 1165
|
|
|
```

RESULT 14

```
A20016
LOCUS A20016 1512 bp DNA linear PAT 14-JUL-1995
DEFINITION SEQ ID NO: 24; Nucleotide sequence for streptokinase fused to yeast
alpha-factor.
ACCESSION A20016
VERSION A20016.1 GI:1247850
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 1512)
AUTHORS
TITLE PROTEINS AND NUCLEIC ACIDS
JOURNAL Patent: WO 9109125-A 24 27-JUN-1991;
FEATURES
source
1. 1512
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
7. 1506
/codon_start=1
/transl_table=11
/product="streptokinase fused to yeast alpha-factor"
/protein_id="CAA01487.1"
/db_xref="GI:1247851"
/translation="MRFP81FTAVLFAASSALAAAPVNTTTEDETAQIPAEAVIGYLDL
EGDPDVALPFSNSTNGLLFNTTITASIATKEGVSLDKRIAGPMLDRPSVNNQ
LVSVAGTVEGTNQDLSLKEPFDLTSRPAHGKTEQGLSPKSPKATPSGAMPKLE
KADLLKALQESLIANVSHNSDDYFEVIDFASDATITDRNGKVIFADKQSVLTPQVQ
BFLSGHVRVPRYPKSPKIQNAKSVYVERDSSIVPLNDNDIPRILPMDQSFVHKVRE
DTITSQELLAAQASINLTKHPGTIVYKVKKEPKYDPDRSHLKLPTIKYDVNTNEL
QAYEINKSGNEELNTHLISEKYYLKKKEPKYDPDRSHLKLPTIKYDVNTNEL
LKSEOLLTASRENLDPRDLYDRDKAKLLYNLDAFGINDYTLTGKVEDNHDNTNRII
TVMGKREGENASHTLAYDKDRYTEEREVSYLYRTGTPIDPNFNDK"
```

ORIGIN

```
Query Match 72.0%; Score 1110.2; DB 6; Length 1512;
Best Local Similarity 97.2%; Pred. No. 1.7e-272;
Matches 1130; Conservative 0; Mismatches 33; Indels 0; Gaps 0;
QY 25 ATACCATGATAGTGGTCTGGAATGGCTACTAGATCGTCTTCTGTGTAAATACAGCCAAAT 84
|
|
|
Db 254 ATAAAGAAATTCCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAAAT 313
|
|
|
QY 85 TGGTTGTTAGGTTGCTGGTACTCTTGGAGGGGAGCAATCAAGACATTTAGTCTTAAATTTT 144
|
|
|
Db 314 TAGTTGTTAGCGTTGCTGGTACTCTTGGAGGGGAGCAATCAAGACATTTAGTCTTAAATTTT 373
|
|
|
QY 145 TTGAAATCGATTAACATACGACCTGCTCATGAGGAAAGACAGACGAGCGTTTAAATGTC 204
|
|
|
Db 374 TTGAAATGACCTAACATACGACCTGCTCATGAGGAAAGACAGACGAGCGTTTAAATGTC 433
|
|
|
QY 205 CAAATCAAAACCAATTTGCTACTGATAGTGGCGGATGTACATAAACTTGAGAAAGCTG 264
|
|
|
Db 434 CAAATCAAAACCAATTTGCTACTGATAGTGGCGGATGTACATAAACTTGAGAAAGCTG 493
|
|
|
QY 265 ACTTACTAAAGGCTATTCAAGAACCAATTGATCGCTTAAGTCCACAGTAAACGACGACTACT 324
|
|
|
Db 494 ACTTACTAAAGGCTATTCAAGAACCAATTGATCGCTTAAGTCCACAGTAAACGACGACTACT 553
|
|
|
QY 325 TTGAGGTCATTTGATTTTGCAGCGCATGCAACCATTTACTGATCGGAAACGCGAAGGCTTACT 384
|
|
|
Db 554 TTGAGGTCATTTGATTTTGCAGCGCATGCAACCATTTACTGATCGGAAACGCGAAGGCTTACT 613
|
|
|
QY 385 TTGCTGACAAAGATGGTTTCGGTAACTTTGCGGACCCCAACCTGTCCAGAAATTTTGTCTAA 444
|
|
|
Db 614 TTGCTGACAAAGATGGTTTCGGTAACTTTGCGGACCCCAACCTGTCCAGAAATTTTGTCTAA 673
|
|
|
QY 445 GCGGACATGTGCGGTTAGACCATATATAAGAAAAACCAATACAAACCAAGCGAAATCTG 504
|
|
|
Db 674 GCGGACATGTGCGGTTAGACCATATATAAGAAAAACCAATACAAACCAAGCGAAATCTG 733
|
|
|
QY 505 TTGATGTTGGAATATATCTGTACAGTTTACTCCCTTTAAACCCCTGATGACGATTTTCAGACAG 564
|
|
|
Db 734 TTGATGTTGGAATATATCTGTACAGTTTACTCCCTTTAAACCCCTGATGACGATTTTCAGACAG 793
|
|
|
QY 565 GTCTCAAGAGTACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACTCTCAAG 624
|
|
|
Db 794 GTCTCAAGAGTACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACTCTCAAG 853
|
|
|
QY 625 AATTACTAGCTCAAGCACAAGCAATTTTAAACAAAAACCCATCCAGGCTATACGATTTATG 684
|
|
|
Db 854 AATTACTAGCTCAAGCACAAGCAATTTTAAACAAAAACCCATCCAGGCTATACGATTTATG 913
|
|
|
QY 685 AACGTGACTCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTTACCAATGG 744
|
|
|
Db 914 AACGTGACTCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTTACCAATGG 973
|
|
|
QY 745 ATCAAGAGTTTACTTACCGTGTAAAAATCGGGAACAAGCTTATAGGATCAATAAAAAAT 804
|
|
|
Db 974 ATCAAGAGTTTACTTACCATGTCAAAATCGGGAACAAGCTTATAGGATCAATAAAAAAT 1033
|
|
|
QY 805 CTGGTCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTTACGTCCTTA 864
|
|
|
Db 1034 CTGGTCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTTACGTCCTTA 1093
|
|
|
QY 865 AAAAGGGGAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCAACCATCA 924
|
|
|
Db 1094 AAAAGGGGAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCAACCATCA 1153
|
|
|
QY 925 AATACGTTGATGTCGATACCAACGAATTTGCTTAAAAAGTGAAGAGCTCTTAAACAGCTAGCG 984
|
|
|
Db 1154 AATACGTTGATGTCGATACCAACGAATTTGCTTAAAAAGTGAAGAGCTCTTAAACAGCTAGCG 1213
|
|
|
QY 985 AAGCTAACTTACACTTCAGAGATTTTATACGATCTCTGATAGGCTTAACTACTCTACA 1044
|
|
|
Db 1214 AAGCTAACTTACACTTCAGAGATTTTATACGATCTCTGATAGGCTTAACTACTCTACA 1273
|
|
|
```

Qy	1045	ACAATCTCGATGCTTTTGGTATTATGAGCATATACCTTAACTGGAAAAATGATGAGATTAATC	1104
Db	1274	ACAATCTCGATGCTTTTGGTATTATGAGCATATACCTTAACTGGAAAAATGATGAGATTAATC	1333
Qy	1105	ACGATGACACCAACCGTATCATACCGTTTATATGGCGAAGCGACCCGAAGGAGAGAATG	1164
Db	1334	ACGATGACACCAACCGTATCATACCGTTTATATGGCGAAGCGACCCGAAGGAGAGAATG	1393
Qy	1165	CTAGCTACCACTTATAGCTGGTGGT	1187
Db	1394	CTAGCTATCACTTATAGCTATGAT	1416
RESULT 15			
I13204			
LOCUS			
DEFINITION Sequence 27 from patent US 5434073.			
ACCESSION I13204			
VERSION I13204.1 GI:910552			
KEYWORDS			
SOURCE			
ORGANISM			
Unknown.			
Unclassified.			
1 (bases 1 to 1512)			
Dawson,K., Hunter,M.G. and Czaplowski,L.G.			
Fibrinolytic and anti-thrombotic cleavable dimers			
Patent: US 5434073-A 27 18-JUL-1995;			
JOURNAL Location/Qualifiers			
FEATURES			
source			
1..1512			
/organism="unknown"			
/mol_type="unassigned DNA"			
ORIGIN			
Query Match 72.0%; Score 1110.2; DB 6; Length 1512;			
Best Local Similarity 97.2%; Pred. No. 1.7e-272;			
Matches 1130; Conservative 0; Mismatches 33; Indels 0; Gaps 0;			
Qy	25	ATACCATGATAGCTGGTCCGTAATGGCTACTAGATCGTCTCTTCGTGTAATAACAGGCCAAT	84
Db	254	ATAAAAGAAATTGCTGGACCTGAGTGGCTGTCTAGACCGTCCATCTGTCAACACAGCCCAAT	313
Qy	85	TGGTTGTTAGCGTTGCTGGTACTGTTTCAGGGGAGCAATCAAGACATTAGTCTTTAAATTTT	144
Db	314	TAGTTGTTAGCGTTGCTGGTACTGTTGAGGGAGCAATCAAGACATTAGTCTTTAAATTTT	373
Qy	145	TTGAAATCGATCTAACTACACGACCTGCTCATGAGGAGAAAGACAGACCAAGGCTTTAAGTC	204
Db	374	TTGAAATTGACCTTAACATCAGACCTGCTCATGAGGAGAAAGACAGACCAAGGCTTTAAGTC	433
Qy	205	CAAAATCAAAACCAATTTGCTACTGATAGTGGCGGATGTCACATAAACTTTGAGAAAGCTG	264
Db	434	CAAAATCAAAACCAATTTGCTACTGATAGTGGCGGATGCCACATAAACTTTGAAAAAGCTG	493
Qy	265	ACTTACTAAAGGCTATTCAAGAACAAATTGATCGTTCACGTTCCACAGTAAACGAGACTACT	324
Db	494	ACTTACTAAAGGCTATTCAAGAACAAATTGATCGTTCACGTTCCACAGTAAACGAGACTACT	553
Qy	325	TTGAGGCTATTGATTTTTCGAAGCGATGCAACCAATTAATCTGATCGAAACGGCAAGGCTTACT	384
Db	554	TTGAGGCTATTGATTTTTCGAAGCGATGCAACCAATTAATCTGATCGAAACGGCAAGGCTTACT	613
Qy	385	TTGCTGACAAAGATGGTTCGGTAACTTTCGGACCGCAACCTGTCACAGATTTTTTGCTAA	444
Db	614	TTGCTGACAAAGATGGTTCGGTAACTTTCGGACCGCAACCTGTCACAGATTTTTTGCTAA	673
Qy	445	GCGGACATGTGCGGGTTAGACCAATATAAAGAAAAACCAATACAAAAACCAAGCGAAATCTG	504
Db	674	GCGGACATGTGCGGGTTAGACCAATATAAAGAAAAACCAATACAAAAACCAAGCGAAATCTG	733
Qy	505	TTGATGTGGAAATATACGTACGTTTACTCCCTTTAAACCTCGATGACGATTTACAGACCAG	564
Db	734	TTGATGTGGAAATATACGTACGTTTACTCCCTTTAAACCTCGATGACGATTTACAGACCAG	793

565	Qy	GTCTCAAGAGATCTAAGAGTATTGAAACAACCTAGCTATCGGTGACACATCACTCTCAAG	624
566			
567			
568			
569			
570			
571			
572			
573			
574			
575			
576			
577			
578			
579			
580			
581			
582			
583			
584			
585			
586			
587			
588			
589			
590			
591			
592			
593			
594			
595			
596			
597			
598			
599			
600			
601			
602			
603			
604			
605			
606			
607			
608			
609			
610			
611			
612			
613			
614			
615			
616			
617			
618			
619			
620			
621			
622			
623			
624			
625	Qy	AATTACTAGCTCAAGCACAAAGCATTTTAAACAAAACACACCCAGCGCTATACGATTTATG	684
626			
627			
628			
629			
630			
631			
632			
633			
634			
635			
636			
637			
638			
639			
640			
641			
642			
643			
644			
645			
646			
647			
648			
649			
650			
651			
652			
653			
654			
655			
656			
657			
658			
659			
660			
661			
662			
663			
664			
665			
666			
667			
668			
669			
670			
671			
672			
673			
674			
675			
676			
677			
678			
679			
680			
681			
682			
683			
684			
685	Qy	AACGTCAGCTCCTCAATCGTCACCTCATGACAATGACATTTTCGTACGATTTTACCAATGG	744
686			
687			
688			
689			
690			
691			
692			
693			
694			
695			
696			
697			
698			
699			
700			
701			
702			
703			
704			
705			
706			
707			
708			
709			
710			
711			
712			

Search completed: February 1, 2006, 04:17:46
Job time : 7836.07 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2006, 10:44:08 ; Search time 931.565 Seconds
(without alignments)
11024.772 Million cell updates/sec

Title: US-09-940-235-9
Perfect score: 1541
Sequence: 1 ttgttgaactttaagaagg.....ccttcacgatgttgtag 1541

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 4996997 seqs, 332346308 residues
Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_21.*
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004as.*
13: Geneseqn2004bs.*
14: Geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1541	100.0	1541	3	Aaa37644 Chimeric
2	1496.2	97.1	2096	3	Aaa37643 Chimeric
3	1185	76.9	1661	3	Aaa37637 Chimeric
4	1179	76.5	1327	3	Aaa37628 Streptoki
5	1166.2	75.7	1377	3	Aaa37622 Streptoki
6	1134.2	73.6	1782	3	Aaa37642 Chimeric
7	1134	73.6	2385	2	Aax80497 Streptoki
8	1132.8	73.5	2030	2	Aaql1651 FB-FB-SK
9	1132.6	73.5	1242	2	Aax80492 Streptoco
10	1132.6	73.5	1245	3	Aaa37633 Streptoki
11	1132.6	73.5	1254	6	Abao5546 Streptoki
12	1132.6	73.5	8893	6	Abao5547 Streptoco
13	1131	73.4	1242	2	Aax16632 Streptoco
14	1127.8	73.2	1242	2	Aax16633 Streptoco
15	1120.2	72.7	7057	12	Adm01294 Plasmid p
16	1118.6	72.6	2566	2	Aat77778 Coding se
17	1110.2	72.0	1407	1	Aan70106 DNA encod
18	1110.2	72.0	1512	2	Aaql2158 Streptoki
19	1108.6	71.9	1245	2	Aaq20665 SKC-2 str

20	1108.6	71.9	1335	2	AAQ12156 Streptoki
21	1108.6	71.9	1458	2	AAQ12162 Factor Xa
22	1108.6	71.9	1467	2	AAQ12490 Factor Xa
23	1108.6	71.9	2589	2	AAQ12160 OmpAL str
24	1100.6	71.4	1473	2	AAQ05603 Sequence
25	1094.2	71.0	2568	1	AAN50493 Streptoki
26	1092.6	70.9	1323	2	AAT29961 Vector ps
27	1084.4	70.4	1209	3	AAX99249 DNA encod
28	1084.4	70.4	1245	3	AAX99250 DNA encod
29	1083.8	70.3	1122	3	AAX99251 DNA encod
30	1083.8	70.3	1158	3	AAX99252 DNA encod
31	1083.4	70.3	1242	5	Aaf82144 Mutant st
32	1078.8	70.0	2253	2	AAQ12161 Met-core
33	1077.6	69.9	1119	2	AAQ12159 Truncated
34	993.4	64.5	1320	6	ABN70192 Streptoco
35	993.4	64.5	1323	13	ADR83811 S. pyogen
36	988.6	64.2	1473	2	AAQ05604 Streptoki
37	973	63.1	1068	2	AAX80493 Recombina
38	968.4	62.8	2208	2	AAX83589 Recombina
39	894.6	58.1	1245	10	ADP48644 Streptoco
40	761.2	49.4	1262	2	AAQ10230 Synthetic
41	450	29.2	450	2	AAX80494 Recombina
42	330	21.4	777	3	AAA37632 Human fib
43	330	21.4	1929	13	AD817488 Nucleotid
44	330	21.4	1929	13	ADR97657 Human fib
45	330	21.4	1929	14	ADW44479 Human fib

ALIGNMENTS

RESULT 1
AAA37644
ID AAA37644 standard; DNA; 1541 BP.
XX
AC AAA37644;
XX
15-SEP-2003 (revised)
DT 13-OCT-2000 (first entry)
XX
Chimeric SK-FBD coding sequence.
XX
Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
KW plasminogen; human; fibronectin; thrombolytic therapy;
KW cardiovascular disorder; ss.
XX
Streptococcus dysgalactiae subsp. equisimilis.
OS Homo sapiens.
OS Chimeric.
XX
EP1024192-A2.
XX
02-AUG-2000.
XX
23-DEC-1999; 99EP-00310541.
XX
24-DEC-1998; 98TN-DE003825.
XX
(COUL) CSIR COUNCIL SCI IND RES.
XX
Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
Yadav M;
XX
WPI; 2000-516032/47.
XX
Hybrid streptokinase-fibrin binding domain polypeptides useful for
thrombolytic therapy comprises a streptokinase fused with fibrin binding
domains of human fibronectin.
XX
Disclosure; Fig 17b; 58pp; English.
XX
This sequence represents a chimeric streptokinase-fibrin binding domain
(SK-FBD) protein coding sequence. The invention relates to a hybrid

CC plasminogen activator (PA) comprises a polypeptide fusion between
 CC streptokinase (SK), which are capable of plasminogen (PG) activation, and
 CC fibrin binding regions of human fibrinectin, which are from fibrin
 CC binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA possesses the
 CC ability to bind with fibrin independently and also characteristically
 CC retains a PG activation ability which becomes evident only after a
 CC pronounced duration, or lag, after exposure of the PA to a suitable
 CC animal or human PG. The hybrid streptokinase-fibrin binding domain
 CC polypeptides are useful in thrombolytic therapy for various kinds of
 CC cardiovascular disorders. The hybrids have enhanced fibrin selectivity as
 CC well as kinetics of plasminogen activation that are distinct from that of
 CC natural streptokinase in being characterised by a temporary delay, or lag
 CC of several minutes in the natural rate of the catalytic conversion of
 CC plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins
 CC can bind tightly with fibrin in blood clots soon after introduction into
 CC the vascular system without significantly activating the circulating
 CC blood plasminogen to plasmin, thus aiding in the localisation of the
 CC plasminogen activation process to the site of pathological thrombus. This
 CC overcomes systemic plasminogen activation encountered during clinical use
 CC of streptokinase. (Updated on 15-SEP-2003 to standardise OS field)
 XX
 SQ Sequence 1541 BP; 497 A; 328 C; 335 G; 381 T; 0 U; 0 Other;

Query Match 100.0%; Score 1541; DB 3; Length 1541;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1541; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTTGTTTAACTTAAGAGGAGATATACCATGATAGCTGGTCTCGAATGGCTACTAGATC 60
 Db |||||
 Qy 1 TTTGTTTAACTTAAGAGGAGATATACCATGATAGCTGGTCTCGAATGGCTACTAGATC 60
 Db |||||
 Qy 61 GTCTCTCTGTAATAACGCCAATTTGTTAGCGTTGCTGTTAGTGGTCTGTTAGGGGACGA 120
 Db |||||
 Qy 61 GTCTCTCTGTAATAACGCCAATTTGTTAGCGTTGCTGTTAGTGGTCTGTTAGGGGACGA 120
 Db |||||
 Qy 121 ATCAAGACATTAGTCTTAAATTTTGAATCGATCTAATCAACGACCTGCTCATGGAG 180
 Db |||||
 Qy 121 ATCAAGACATTAGTCTTAAATTTTGAATCGATCTAATCAACGACCTGCTCATGGAG 180
 Db |||||
 Qy 181 GAAAGACAGAGGAGCTTAAGTCAAAATCAAAACCAATTTGCTACTGATAGTGGCGCA 240
 Db |||||
 Qy 181 GAAAGACAGAGGAGCTTAAGTCAAAATCAAAACCAATTTGCTACTGATAGTGGCGCA 240
 Db |||||
 Qy 241 TGTCATAAATTTGAGAAAGCTGACTTAAAGCTATTCAAGACCAATTTGATCGCTA 300
 Db |||||
 Qy 241 TGTCATAAATTTGAGAAAGCTGACTTAAAGCTATTCAAGACCAATTTGATCGCTA 300
 Db |||||
 Qy 301 ACGTCCACAGTAAACGACGACTTCTGAGGTCAATTTGCAAGCGATGCAACCATTA 360
 Db |||||
 Qy 301 ACGTCCACAGTAAACGACGACTTCTGAGGTCAATTTGCAAGCGATGCAACCATTA 360
 Db |||||
 Qy 361 CTGATCGAAAACGGCAAGGCTCTACTTTGCTGCAAAAGATGGTTTCGGTAACTTCCGACCC 420
 Db |||||
 Qy 361 CTGATCGAAAACGGCAAGGCTCTACTTTGCTGCAAAAGATGGTTTCGGTAACTTCCGACCC 420
 Db |||||
 Qy 421 AACCTGTCCAGAAATTTTGTCTAAGCGGACATGTGGCGTTAGACCATATAAGAAAAAC 480
 Db |||||
 Qy 421 AACCTGTCCAGAAATTTTGTCTAAGCGGACATGTGGCGTTAGACCATATAAGAAAAAC 480
 Db |||||
 Qy 481 CAATACAAAAACCAAGCGAAATCTGTTGATGTGGAATATATCTGTACAGTTTACTCCCTTAA 540
 Db |||||
 Qy 481 CAATACAAAAACCAAGCGAAATCTGTTGATGTGGAATATATCTGTACAGTTTACTCCCTTAA 540
 Db |||||
 Qy 541 ACCCTGATGACGATTTTCAGACAGGTTCTCAAGATCTAAGCTATTGAAAAACACTAGCTA 600
 Db |||||
 Qy 541 ACCCTGATGACGATTTTCAGACAGGTTCTCAAGATCTAAGCTATTGAAAAACACTAGCTA 600
 Db |||||
 Qy 601 TCGGTGACACCATCATCTCAAGAAATTTACTAGCTCAAGCAAAAGCATTTTAAACAAA 660
 Db |||||
 Qy 601 TCGGTGACACCATCATCTCAAGAAATTTACTAGCTCAAGCAAAAGCATTTTAAACAAA 660
 Db |||||
 Qy 661 ACCACCCAGGCTATACGATTTATGACGCTGCTCTCAATCTGCTCATGATGACATGACA 720
 Db |||||

Db 661 ACCACCCAGGCTATACGATTTTAAAGCGTGACTCTCTCAATCGTCACTCATGACAAATGACA 720
 Qy |||||
 Qy 721 TTTTCCGTACGATTTTACCAATCGATCAAGAGTTTACTTACCGTGTAAATAATCGGAAC 780
 Db |||||
 Db 721 TTTTCCGTACGATTTTACCAATCGATCAAGAGTTTACTTACCGTGTAAATAATCGGAAC 780
 Qy |||||
 Qy 781 AAGCTTATAGGATCAATAAAAAATCTGGTCTGAATGAAGAAATAAAACAACTGACCTGA 840
 Db |||||
 Db 781 AAGCTTATAGGATCAATAAAAAATCTGGTCTGAATGAAGAAATAAAACAACTGACCTGA 840
 Qy |||||
 Qy 841 TCTCTGAGAAATATACGCTCTTAAAAAGGGGAAAGCCGTATGATCCCTTTGATCGCA 900
 Db |||||
 Db 841 TCTCTGAGAAATATACGCTCTTAAAAAGGGGAAAGCCGTATGATCCCTTTGATCGCA 900
 Qy |||||
 Qy 901 GTCACCTGAAACTGTTTACCATCAATACGTTTATGATGTCGATACCAACGAATTTGCTAAAA 960
 Db |||||
 Db 901 GTCACCTGAAACTGTTTACCATCAATACGTTTATGATGTCGATACCAACGAATTTGCTAAAA 960
 Qy |||||
 Qy 961 GTGAGCAGCTCTTAAACAGCTAGCGAAAGCTAACTTACAGCTTTCAGAGATTTATACGATCCTC 1020
 Db |||||
 Db 961 GTGAGCAGCTCTTAAACAGCTAGCGAAAGCTAACTTACAGCTTTCAGAGATTTATACGATCCTC 1020
 Qy |||||
 Qy 1021 GTGATAAGGCTAACTACTCTACAACTCTCGATGCTTTTGGTATTATGAGACTATACCT 1080
 Db |||||
 Db 1021 GTGATAAGGCTAACTACTCTACAACTCTCGATGCTTTTGGTATTATGAGACTATACCT 1080
 Qy |||||
 Qy 1081 TAACTGAAAAAGTAGAGGATATACGATGACACCAACCGTATCATACCGTTTATATGG 1140
 Db |||||
 Db 1081 TAACTGAAAAAGTAGAGGATATACGATGACACCAACCGTATCATACCGTTTATATGG 1140
 Qy |||||
 Qy 1141 GCAAGCGACCCGAAAGGAGAGAAATGCTAGCTTACCATTTAGCTGGTGGCGACGCGCAAC 1200
 Db |||||
 Db 1141 GCAAGCGACCCGAAAGGAGAGAAATGCTAGCTTACCATTTAGCTGGTGGCGACGCGCAAC 1200
 Qy |||||
 Qy 1201 AGATTGTACCCATAGCTGAGAAAGTGTGTTTGTATCATGCTGCTGGGACTTCTCTATGTGTCG 1260
 Db |||||
 Db 1201 AGATTGTACCCATAGCTGAGAAAGTGTGTTTGTATCATGCTGCTGGGACTTCTCTATGTGTCG 1260
 Qy |||||
 Qy 1261 GAGAAAACGTGGGAGAAAGCCCTACCAAGGCTGGATGATGTTGATGTTGTTGCTGGGAG 1320
 Db |||||
 Db 1261 GAGAAAACGTGGGAGAAAGCCCTACCAAGGCTGGATGATGTTGATGTTGTTGCTGGGAG 1320
 Qy |||||
 Qy 1321 AAGCGACGCGACGCACTTGCATCTTAGAATAGATGCAACGATCAGGACACAGGA 1380
 Db |||||
 Db 1321 AAGCGACGCGACGCACTTGCATCTTAGAATAGATGCAACGATCAGGACACAGGA 1380
 Qy |||||
 Qy 1381 CATCTTATAGAAATTGGAGACACCTCGAGCAAGAGGATATCGAGGAAACCTGCTCCAGT 1440
 Db |||||
 Db 1381 CATCTTATAGAAATTGGAGACACCTCGAGCAAGAGGATATCGAGGAAACCTGCTCCAGT 1440
 Qy |||||
 Qy 1441 GCATCTGCAAGGCAACCGCCGAGGAGAGTGGAAAGTGTGAGAGGCAACCTCTGTGCGAGA 1500
 Db |||||
 Db 1441 GCATCTGCAAGGCAACCGCCGAGGAGAGTGGAAAGTGTGAGAGGCAACCTCTGTGCGAGA 1500
 Qy |||||
 Qy 1501 CCACATCGAGCGGATCTGCGCCCTTCCACGATGTTCTGTTAG 1541
 Db |||||
 Db 1501 CCACATCGAGCGGATCTGCGCCCTTCCACGATGTTCTGTTAG 1541

RESULT 2

AAA37643

ID AAA37643 standard; DNA; 2096 BP.

XX AAA37643;

XX AC 15-SEP-2003 (revised)

DT 13-OCT-2000 (first entry)

XX Chimeric SK-FBD coding sequence.

XX Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
 KW plasminogen; human; fibrinectin; thrombolytic therapy;
 KW cardiovascular disorder; ss.

XX		
OS	Streptococcus dysgalactiae subsp. equisimilis.	
OS	Homo sapiens.	
XX	Chimeric.	
PN	EP1024192-A2.	
XX		
FD	02-AUG-2000.	
XX		
PF	23-DEC-1999; 99EP-00310541.	
XX		
PR	24-DEC-1998; 96IN-DE003825.	
XX		
PA	(COUL) CSIR COUNCIL SCI IND RES.	
XX		
PI	Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;	
PI	Yadav M;	
XX		
DR	WPI; 2000-516032/47.	
XX		
PT	Hybrid streptokinase-fibrin binding domain polypeptides useful for	
PT	thrombolytic therapy comprises a streptokinase fused with fibrin binding	
PT	domains of human fibronectin.	
XX		
PS	Example 6; Fig 22b; 58pp; English.	
XX		
CC	This sequence represents a chimeric streptokinase-fibrin binding domain	
CC	(SK-FBD) protein coding sequence. The invention relates to a hybrid	
CC	plasminogen activator (PA) comprises a polypeptide fusion between	
CC	streptokinase (SK), which are capable of plasminogen (PG) activation, and	
CC	fibrin binding regions of human fibronectin, which are from fibrin	
CC	binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA possesses the	
CC	ability to bind with fibrin independently and also characteristically	
CC	retains a PG activation ability which becomes evident only after a	
CC	prounced duration, or lag, after exposure of the PA to a suitable	
CC	animal or human PG. The hybrid streptokinase-fibrin binding domain	
CC	polypeptides' are useful in thrombolytic therapy for various kinds of	
CC	cardiovascular disorders. The hybrids have enhanced fibrin selectivity as	
CC	well as kinetics of plasminogen activation that are distinct from that of	
CC	natural streptokinase in being characterised by a temporary delay, or lag	
CC	of several minutes in the natural rate of the catalytic conversion of	
CC	plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins	
CC	can bind tightly with fibrin in blood clots soon after introduction into	
CC	the vascular system without significantly activating the circulating	
CC	blood plasminogen to plasmin, thus aiding in the localisation of the	
CC	plasminogen activation process to the site of pathological thrombus. This	
CC	overcomes systemic plasminogen activation encountered during clinical use	
CC	of streptokinase. (Updated on 15-SEP-2003 to standardise OS field)	
XX		
SQ	Sequence 2096 BP; 643 A; 480 C; 484 G; 489 T; 0 U; 0 Other;	
	Query Match 97.1%; Score 1496.2; DB 3; Length 2096;	
	Best Local Similarity 99.5%; Pred. No. 0;	
	Matches 1501; Conservative 0; Mismatches 8; Indels 0; Gaps 0	
Qy	33 ATAGCTGTCCTCGAATGGCTACTAGATCGTCTCTTGTAATAACAGCCAATTGGTTGT 92	
Db	588 ATTGCTGGACCTGAGTGCTGTAGACCGTCCATCTGTCAACAACAGCCAATTGGTTGT 647	
Qy	93 AGCGTTGCTGGTACATGTTGAGGGGACGAATCAAGAATCATTAGTCTTTAAATTTTTGAAATC 152	
Db	648 AGCGTTGCTGGTACTGTTTGAGGGGACGAATCAAGACATTAGTCTTTAAATTTTTGAAATC 707	
Qy	153 GATCTTAACATCACGACCTGCTCATNGAGGGAAGACAGACGAGCGTTAAGTCCAAATCA 212	
Db	708 GATCTTAACATCACGACCTGCTCATNGAGGGAAGACAGACGAGCGTTAAGTCCAAATCA 767	
Qy	213 AAACCAATTTGCTACTGATAGTGGGCCGATGTCAACATAACTTGAGAAAGCTGCATTACTA 272	
Db	768 AAACCAATTTGCTACTGATAGTGGGCCGATGTCAACATAACTTGAGAAAGCTGCATTACTA 827	
Qy	273 AAGGCTATTCAAGAACAAATGATCGCTTAACGCTCCACAGTAACGACGACTCTTTGAGGTC 332	

Db	828	AAGGCTATTCAAGAACAAATTGATCGCTAAACGTCACAGTACCGACGACTACTTTGAGGTC	887
Qy	333	ATTGATTTTGCACGCGATGCAACCAATTACTGATCGAAACGCAAGGTCTACTTTGCTGTAC	392
Db	888	ATTGATTTTGCACGCGATGCAACCAATTACTGATCGAAACGCAAGGTCTACTTTGCTGTAC	947
Qy	393	AAAGATGGTTCCGGTAAACCTTGCCGCAACCACTGTCCTCAAGAAATTTTGTCTAAGCGGCAAT	452
Db	948	AAAGATGGTTCCGGTAAACCTTGCCGCAACCACTGTCCTCAAGAAATTTTGTCTAAGCGGCAAT	1007
Qy	453	GTGCGCGTTAGACCAATATAAGAAACCAATATACAAACCAAGCGAAATCTGTTGATGTG	512
Db	1008	GTGCGCGTTAGACCAATATAAGAAACCAATATACAAACCAAGCGAAATCTGTTGATGTG	1067
Qy	513	GAATATACTGTACAGTTTACTTCCCTTAAACCCCTGATGACGANTTTCAGACAGGTCCTCAAA	572
Db	1068	GAATATACTGTACAGTTTACTTCCCTTAAACCCCTGATGACGANTTTCAGACAGGTCCTCAAA	1127
Qy	573	GATCTAAGCTATTGAAACCACTAGCTPATCGGTGACACCACTCATCTCAAGAATTTACTTA	632
Db	1128	GATCTAAGCTATTGAAACCACTAGCTPATCGGTGACACCACTCATCTCAAGAATTTACTTA	1187
Qy	633	GCTCAAGCACAAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTTATGAACGTGAC	692
Db	1188	GCTCAAGCACAAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTTATGAACGTGAC	1247
Qy	693	TCCTCAATCGTCACCTCATGACAAATGACATTTTCGGTACGATTTTACCAATGGATCAAGAG	752
Db	1248	TCCTCAATCGTCACCTCATGACAAATGACATTTTCGGTACGATTTTACCAATGGATCAAGAG	1307
Qy	753	TTTACTTACCGTGTTTAAAAATCGGGAAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG	812
Db	1308	TTTACTTACCGTGTTTAAAAATCGGGAAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG	1367
Qy	813	AATGAAGAATAAACAACACTGCACTCTCTGAGAAATATTACGTCCTTAAAAAGGG	872
Db	1368	AATGAAGAATAAACAACACTGCACTCTCTGAGAAATATTACGTCCTTAAAAAGGG	1427
Qy	873	GAAAAGCGTATGATCCCTTTGATCGCAGTCACCTTGAAACGTGTTACCATCAAAATAGTT	932
Db	1428	GAAAAGCGTATGATCCCTTTGATCGCAGTCACCTTGAAACGTGTTACCATCAAAATAGTT	1487
Qy	933	GATGTCGATACCAACGAATTTCTAAAAAGTGAGCAGCTCTTAAACAGCTACGCAACGTAAC	992
Db	1488	GATGTCGATACCAACGAATTTCTAAAAAGTGAGCAGCTCTTAAACAGCTACGCAACGTAAC	1547
Qy	993	TTAGACTTCAGAGATTTTATAGATCTCTGCTGTGATAAGGCTAAACTCTCTACAACAATCTC	1052
Db	1548	TTAGACTTCAGAGATTTTATAGATCTCTGCTGTGATAAGGCTAAACTCTCTACAACAATCTC	1607
Qy	1053	GATGCTTTTGATTTATGGAATACCTTAACTGGAAGAGTAGAGGATATACGATGAC	1112
Db	1608	GATGCTTTTGATTTATGGAATACCTTAACTGGAAGAGTAGAGGATATACGATGAC	1667
Qy	1113	ACCAACCGTATCATAAACCGTTTATATGGGCAAGCGACCCGAAGGAGAGATGCTAGCTAC	1172
Db	1668	ACCAACCGTATCATAAACCGTTTATATGGGCAAGCGACCCGAAGGAGAGATGCTAGCTAC	1727
Qy	1173	CATTTAGCTGGTGGCGCAGGCGCAACAGATTTGATCCCATAGCTGAGAGAGTGTTTGAT	1232
Db	1728	CATTTAGCTGGTGGCGCAGGCGCAACAGATTTGATCCCATAGCTGAGAGAGTGTTTGAT	1787
Qy	1233	CATGCTCTGGGACTTCTTATGTGTCGGAGAACGTGGGAGAGGCCCTACCAAGGCTGG	1292
Db	1788	CATGCTCTGGGACTTCTTATGTGTCGGAGAACGTGGGAGAGGCCCTACCAAGGCTGG	1847
Qy	1293	ATGATGTGTAGATTGTACTTCCCTGGGAGAGGCGAGCGCATCACTTGCACTTCTAGA	1352
Db	1848	ATGATGTGTAGATTGTACTTCCCTGGGAGAGGCGAGCGCATCACTTGCACTTCTAGA	1907
Qy	1353	AATAGATGCAACGATCAGGACACAAGGACATCTTATAGAAATCGGAGACCTCGAGCAAG	1412
Db	1908	AATAGATGCAACGATCAGGACACAAGGACATCTTATAGAAATCGGAGACCTCGAGCAAG	1967

Db 891 TCTCTGAGAAATATTACGTCCTTAAAAAAGGGGAAAAAGCGTATGATCCCTTTGATCGCA 950
 Qy 901 GTCACCTTGAACCTGTTTACCACCAATACGTTGATGTCGATACCAACGAATGCTAATA 960
 Db 951 GTCACCTTGAACCTGTTTACCACCAATACGTTGATGTCGATACCAACGAATGCTAATA 1010
 Qy 961 GTGAGCAGCTCTTAAACAGCTAGCGAACGTAACCTTAGACTTCAGAGATTTATACGATCCCTC 1020
 Db 1011 GTGAGCAGCTCTTAAACAGCTAGCGAACGTAACCTTAGACTTCAGAGATTTATACGATCCCTC 1070
 Qy 1021 GTGATAAGGCTAAACCTTACTCTTACAAACATCTCGATCTTTTGGTATTTATGACCTATACCT 1080
 Db 1071 GTGATAAGGCTAAACCTTACTCTTACAAACATCTCGATCTTTTGGTATTTATGACCTATACCT 1130
 Qy 1081 TAACTGGAAGAGTAGAGGATTAATCAGATGACACCAACCGTATACAAACGGTTTATATGG 1140
 Db 1131 TAACTGGAAGAGTAGAGGATTAATCAGATGACACCAACCGTATACAAACGGTTTATATGG 1190
 Qy 1141 GCAAGCGACCCGAAGGAGAGAGTGTAGCTTACCATTTAGCTGTGTGT 1187
 Db 1191 GCAAGCGACCCGAAGGAGAGAGTGTAGCTTACCATTTAGCTGTGTGT 1237

RESULT 5

AAA37622
 ID AAA37622 standard; DNA; 1377 BP.

XX AAA37622;

AC 15-SEP-2003 (revised)

DT 13-OCT-2000 (first entry)

XX Streptokinase-NTRN gene.

XX Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;

KW plasminogen; human; fibronectin; thrombolytic therapy;

KW cardiovascular disorder; ss.

XX Streptococcus dysgalactiae subsp. equisimilis.

XX EP1024192-A2.

XX 02-AUG-2000.

XX 23-DEC-1999; 99EP-00310541.

XX 24-DEC-1998; 98IN-DE003825.

XX (COUL) CSIR COUNCIL SCI IND RES.

XX Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;

PI Yadav M;

XX WPI; 2000-516032/47.

XX Hybrid streptokinase-fibrin binding domain polypeptides useful for

PT thrombolytic therapy comprises a streptokinase fused with fibrin binding

PT domains of human fibronectin.

XX Example 1; Fig 11; 58pp; English.

XX This sequence represents a streptokinase-NTRN (SK-NTRN) gene (where NTRN

CC stands for N-terminally repaired with native sequence). The invention

CC relates to a hybrid plasminogen activator (PA) comprises a polypeptide

CC fusion between streptokinase (SK), which are capable of plasminogen (PG)

CC activation, and fibrin binding regions of human fibronectin, which are

CC from fibrin binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA

CC possesses the ability to bind with fibrin independently and also

CC characteristically retains a PG activation ability which becomes evident

CC only after a pronounced duration, or lag, after exposure of the PA to a

CC suitable animal or human PG. The hybrid streptokinase-fibrin binding

CC domain polypeptides are useful in thrombolytic therapy for various kinds

CC

CC of cardiovascular disorders. The hybrids have enhanced fibrin selectivity
 CC as well as kinetics of plasminogen activation that are distinct from that
 CC of natural streptokinase in being characterised by a temporary delay, or
 CC lag of several minutes in the natural rate of the catalytic conversion of
 CC plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins
 CC can bind tightly with fibrin in blood clots soon after introduction into
 CC the vascular system without significantly activating the circulating
 CC blood plasminogen to plasmin, thus aiding in the localisation of the
 CC plasminogen activation process to the site of pathological thrombus. This
 CC overcomes systemic plasminogen activation encountered during clinical use
 CC of streptokinase. (Updated on 15-SEP-2003 to standardise OS field)
 XX

SQ Sequence 1377 BP; 465 A; 302 C; 265 G; 345 T; 0 U; 0 Other;

Query Match 75.7%; Score 1166.2; DB 3; Length 1377;

Best Local Similarity 98.9%; Pred. No. 0;

Matches 1174; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 1 TTTGTTTAACTTTAAGAGGAGATATACCATGATAGTGGTCTCTGAATGGCTACTAGATC 60

Db 101 TTTGTTTAACTTTAAGAGGAGATATACCATGATAGTGGTCTCTGAATGGCTACTAGATC 160

Qy 61 GTCCTTCTGTAATAACAGCCAAATGGTTGTTAGCGTTGCTGTACTGTGTGAGGGACGA 120

Db 161 GTCCTTCTGTAATAACAGCCAAATGGTTGTTAGCGTTGCTGTACTGTGTGAGGGACGA 220

Qy 121 ATCAAGACATTAGTCTTAAATTTTGGAAATCGATCTAATCAATCAACGACCTGCTCATGAG 180

Db 221 ATCAAGACATTAGTCTTAAATTTTGGAAATCGATCTAATCAATCAACGACCTGCTCATGAG 280

Qy 181 GAAAGACAGAGCAAGCTTAAAGTCCAAATCAAAATCCATTTGCTGTACTGTAGTGGCGCGA 240

Db 281 GAAAGACAGAGCAAGCTTAAAGTCCAAATCAAAATCCATTTGCTGTACTGTAGTGGCGCGA 340

Qy 241 TGTCACATAAACTTTGAGAAAGCTGACTTACTAAAGGCTATTCAAGACAAATTTGATCGCTA 300

Db 341 TGTCACATAAACTTTGAGAAAGCTGACTTACTAAAGGCTATTCAAGACAAATTTGATCGCTA 400

Qy 301 ACGTCCACAGTAACGACGACTACTTTGAGGTCAATTTGCAAGCGATGCAACCAATTA 360

Db 401 ACGTCCACAGTAACGACGACTACTTTGAGGTCAATTTGCAAGCGATGCAACCAATTA 460

Qy 361 CTGATCGAAACGCGAAGGCTCTCTTTGCTGACAAAGATGGTTCGGTAACCTTCCCGACCC 420

Db 461 CTGATCGAAACGCGAAGGCTCTCTTTGCTGACAAAGATGGTTCGGTAACCTTCCCGACCC 520

Qy 421 AACCTGTCCAAGAAATTTTGTCTAAGCGGACATGTGCGGTTAGACCATATAAAGAAAAAC 480

Db 521 AACCTGTCCAAGAAATTTTGTCTAAGCGGACATGTGCGGTTAGACCATATAAAGAAAAAC 580

Qy 481 CAATACAAAACCAAGCGAAATCTGTTGATGTGGAATATATCTGTACAGTTTACTCCCTTAA 540

Db 581 CAATACAAAACCAAGCGAAATCTGTTGATGTGGAATATATCTGTACAGTTTACTCCCTTAA 640

Qy 541 ACCCTGATGACGATTTTACCAATGATCAAGAGTTTACTTACCGTGTAAAAATCGGGAAC 600

Db 641 ACCCTGATGACGATTTTACCAATGATCAAGAGTTTACTTACCGTGTAAAAATCGGGAAC 700

Qy 601 TCGGTGACACCATCAATCTCAAGAAATTTACTAGCTCAAGCAACAAAGCAATTTTAAACAAA 660

Db 701 TCGGTGACACCATCAATCTCAAGAAATTTACTAGCTCAAGCAACAAAGCAATTTTAAACAAA 760

Qy 661 ACCACGAGCTATACGATTTATGAACGTCCTTCAATCGTCACTCATGACAAATGACA 720

Db 761 ACCACGAGCTATACGATTTATGAACGTCCTTCAATCGTCACTCATGACAAATGACA 820

Qy 721 TTTTCCGTACGATTTTACCAATGATCAAGAGTTTACTTACCGTGTAAAAATCGGGAAC 780

Db 821 TTTTCCGTACGATTTTACCAATGATCAAGAGTTTACTTACCGTGTAAAAATCGGGAAC 880

Qy 781 AAGCTTATAGGATCAATAAAAAATCTGGTCTGAATGAAGAAATAAACAACACTGACCTGA 840

Db 881 AAGCTTATAGGATCAATAAAAAATCTGGTCTGAATGAAGAAATAAACAACACTGACCTGA 940

QY 841 TCTCTGAGAAATATTACGTCCTTAAATAAGGGGAAAGCCGTATGATCCCTTTGATCGCA 900
Dd |||||
QY 901 GTCACTTGAACGTTTCCACATCAATACGTTGATGTCGATACCAACGAATTCCTAAAAA 960
Dd |||||
QY 961 GTGAGCAGCTCTTAAACAGCTAGCGAACTTAACTTACAGATTTATACGATCCTC 1020
Dd |||||
QY 1021 GTGATAAGGCTAACTCTTACAACTCTCGATGCTTTTGGTATTATGCACTATACCT 1080
Dd |||||
QY 1081 TAACTGGAAGTAGAGGATAATCAGATGACACCAACCGTATCATAAACCGTTTATATGG 1140
Dd |||||
QY 1141 GCAAGCGACCCGAGGAGAGAACTAGCTACCATTTAGCTGCTGGT 1187
Dd |||||

RESULT 6

AAA37642
ID AAA37642 standard; DNA; 1782 BP.

AC AAA37642;

DT 15-SEP-2003 (revised)
DT 13-OCT-2000 (first entry)

XX Chimeric SK-FBD coding sequence.

XX Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
KW plasminogen; human; fibrinectin; thrombolytic therapy;
KW cardiovascular disorder; ss.

OS Streptococcus dysgalactiae subsp. equisimilis.

OS Homo sapiens.

OS Chimeric.

PN EP1024192-A2.

XX 02-AUG-2000.

XX 23-DEC-1999; 99EP-00310541.

XX 24-DEC-1998; 98IN-DE003825.

XX (COUL) CSIR COUNCIL SCI IND RES.

XX Sahní G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
PI Yadav M;

XX WPI; 2000-516032/47.

XX Hybrid streptokinase-fibrin binding domain polypeptides useful for
PT thrombolytic therapy comprises a streptokinase fused with fibrin binding
PT domains of human fibrinectin.

XX Example 5; Fig 21b; 58pp; English.

XX This sequence represents a chimeric streptokinase-fibrin binding domain
CC (SK-FBD) protein coding sequence. The invention relates to a hybrid
CC plasminogen activator (PA) comprises a polypeptide fusion between
CC streptokinase (SK), which are capable of plasminogen (PG) activation, and
CC fibrin binding regions of human fibrinectin, which are from fibrin
CC binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA possesses the
CC ability to bind with fibrin independently and also characteristically

CC retains a PG activation ability which becomes evident only after a
CC pronounced duration, or lag, after exposure of the PA to a suitable
CC animal or human PG. The hybrid streptokinase-fibrin binding domain
CC polypeptides are useful in thrombolytic therapy for various kinds of
CC cardiovascular disorders. The hybrids have enhanced fibrin selectivity as
CC well as kinetics of plasminogen activation that are distinct from that of
CC natural streptokinase in being characterised by a temporary delay, or lag
CC of several minutes in the natural rate of the catalytic conversion of
CC plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins
CC can bind tightly with fibrin in blood clots soon after introduction into
CC the vascular system without significantly activating the circulating
CC blood plasminogen to plasmin, thus aiding in the localisation of the
CC plasminogen activation process to the site of pathological thrombus. This
CC overcomes systemic plasminogen activation encountered during clinical use
CC of streptokinase. (Updated on 15-SEP-2003 to standardise OS field)

XX Sequence 1782 BP; 573 A; 407 C; 377 G; 425 T; 0 U; 0 Other;

QY Query Match 73.6%; Score 1134.2; DB 3; Length 1782;

Dd Best Local Similarity 98.9%; Pred. No. 0;

Matches 1142; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 33 ATAGCTGGTCTGTAATGGCTACTAGATCGTCCTCTGTAAATAACAGCAATTCGTTGTT 92

Dd 538 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCCAAATTCGTTGTT 597

QY 93 AGCGTTCGTCTGTTAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC 152

Dd 598 AGCGTTCGTCTGTTAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC 657

QY 153 GATCTAACATCAGCACCTGCTCATGAGAGAAAGACAGACGAAGCTTTAAGTCCAAATCA 212

Dd 658 GATCTAACATCAGCACCTGCTCATGAGAGAAAGACAGACGAAGCTTTAAGTCCAAATCA 717

QY 213 AAACATTTGCTACTGATAGTGGCGCGATGTCAATAAACTTGAGAAGCTGACTTACTA 272

Dd 718 AAACATTTGCTACTGATAGTGGCGCGATGTCAATAAACTTGAGAAGCTGACTTACTA 777

QY 273 AAGCTATTCAAGAACAAATTTGATCGCTTAAACGCTTCCACAGTAAAGCTGAGGTC 332

Dd 778 AAGCTATTCAAGAACAAATTTGATCGCTTAAACGCTTCCACAGTAAAGCTGAGGTC 837

QY 333 ATTGATTTTCAAGCGATGCAACCATTTACTGATCGAAACGCAAGCTTCTTTGCTGAC 392

Dd 838 ATTGATTTTCAAGCGATGCAACCATTTACTGATCGAAACGCAAGCTTCTTTGCTGAC 897

QY 393 AAAGATGTTTCGGTAACTTTGCCGACCCCAACCTGTCCAAGAAATTTTGTAAAGCGGACAT 452

Dd 898 AAAGATGTTTCGGTAACTTTGCCGACCCCAACCTGTCCAAGAAATTTTGTAAAGCGGACAT 957

QY 453 GTGCGGTTAGACCATATAAAGAAAAAACCATAAATAAAGCGAAATCTGTTGATGTG 512

Dd 958 GTGCGGTTAGACCATATAAAGAAAAAACCATAAATAAAGCGAAATCTGTTGATGTG 1017

QY 513 GAATATCTCTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTAAA 572

Dd 1018 GAATATCTCTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTAAA 1077

QY 573 GATACTAAGCTTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTTACTA 632

Dd 1078 GATACTAAGCTTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTTACTA 1137

QY 633 GCTCAAGCAAAAGCAATTTTAAACAAAAACCCAGGCTATACGATTTTATGAACGTCAC 692

Dd 1138 GCTCAAGCAAAAGCAATTTTAAACAAAAACCCAGGCTATACGATTTTATGAACGTCAC 1197

QY 693 TCCTCAATCTGCTCATGACCAATGACATTTTCGTACGATTTTACCAATGGATCAAGAG 752

Dd 1198 TCCTCAATCTGCTCATGACCAATGACATTTTCGTACGATTTTACCAATGGATCAAGAG 1257

QY 753 TTTTACTTACCGTGTAAAAAATCGGGAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 812

Dd 1258 TTTTACTTACCGTGTAAAAAATCGGGAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 1317

Db 1861 GAGTTTACTTACCGTGTAAATAATCGGAAACAGCTTTATAGGATCAATAAAAAATCTGGT 1920
QY 810 CTGAATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAA 869
Db 1921 CTGAATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAA 1980
QY 870 GGGGAAAGCCGATGATCCCTTTGATCGGAGTCACCTTGAACTGTTTACCATCAATATAC 929
Db 1981 GGGGAAAGCCGATGATCCCTTTGATCGGAGTCACCTTGAACTGTTTACCATCAATATAC 2040
QY 930 GTTGATGTGATACCAACGAATGCTAAAGTGGCAGCTCTTAACAGCTAGGACGT 989
Db 2041 GTTGATGTGATACCAACGAATGCTAAAGTGGCAGCTCTTAACAGCTAGGACGT 2100
QY 990 AACTTAGACTTCAGAGATTTATACGATCTCTGATGATAAGGCTAAACTTACTCTACAAAT 1049
Db 2101 AACTTAGACTTCAGAGATTTATACGATCTCTGATGATAAGGCTAAACTTACTCTACAAAT 2160
QY 1050 CTCGATGCTTTGGTATTATGGAATATACCTTAACCTGGAAAGTAGAGGATAATCAGAT 1109
Db 2161 CTCGATGCTTTGGTATTATGGAATATACCTTAACCTGGAAAGTAGAGGATAATCAGAT 2220
QY 1110 GACACCAACCGTATCATACCGTTTATATGGGCAAGGACCGAAGAGAGATGCTAGC 1169
Db 2221 GACACCAACCGTATCATACCGTTTATATGGGCAAGGACCGAAGAGAGATGCTAGC 2280
QY 1170 TACCATTTAGCTGGTGT 1187
Db 2281 TATCATTTAGCCTATGAT 2298

RESULT 8
ID AAQ11651 standard; DNA; 2030 BP.
XX AAQ11651;
AC AAQ11651;
DT 08-JUL-1991 (first entry)
XX
DE FB-PB-SK fusion construct.
XX
KW Fibrin-binding protein; fibrinolysis; intravascular thrombi; fibrinogen;
KW streptokinase; fusion protein; ss.
XX
OS Staphylococcus aureus.
XX
FH Key Location/Qualifiers
FT CDS 10..184
FT FT /*tag= a
FT FT /label= FB monomer
FT CDS 185..358
FT FT /*tag= b
FT FT /label= FB monomer
FT CDS 359..1601
FT FT /*tag= c
FT FT /label= streptokinase
XX
PN US5011686-A.
XX
PD 30-APR-1991.
XX
PF 15-NOV-1989; 89US-00437769.
XX
PR 21-SEP-1987; 87US-00099242.
XX
PA (CREA-) CREATIVE BIOMOLEC.
XX
PI Pang RHL;
XX
XX WPI; 1991-140198/19.
DR P-PSDB; AAR11829.
XX
PT Imparting injectable fibrinolytic agent - with affinity for intravascular

PT thrombus, by linking agent to fibrin binding domain.
XX Disclosure; Fig 5; 18pp; English.
XX
CC The DNA encodes an FB-PB dimer linked to the streptokinase coding
CC sequence. The FB fragment has selective affinity for fibrin, low affinity
CC for fibrinogen, and minimal immunogenicity, imparting thrombus-targeting
CC capability. See also AAQ11649 and AAQ11650
XX
SQ Sequence 2030 BP; 673 A; 464 C; 406 G; 487 T; 0 U; 0 Other;
Query Match 73.5%; Score 1132.8; DB 2; Length 2030;
Best Local Similarity 98.5%; Pred. No. 0;
Matches 1143; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
QY 28 CCATGATAGCTGGTCCCTGAATGGCTACTAGATCGTCTTCTGTAATAACAGCAATGG 87
Db 353 CGAAGATTGCTGGACCTGGAGTGGCTGTAGACCGTCCATCTGTCAACACACCAATTAG 412
QY 88 TTGTTAGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGG 147
Db 413 TTGTTAGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGG 472
QY 148 AAATCGATCTAAACATCAACGACCTGCTCATGGAGGAAAGACAGCAAGGCTTAAAGTCAA 207
Db 473 AAATCGATCTAAACATCAACGACCTGCTCATGGAGGAAAGACAGCAAGGCTTAAAGTCAA 532
QY 208 AATCAAAACCAATTTGCTACTGATAGTGGCGGATGTCAATAAATTTGAGAAAGCTGACT 267
Db 533 AATCAAAACCAATTTGCTACTGATAGTGGCGGATGTCAATAAATTTGAGAAAGCTGACT 592
QY 268 TACTAAAGGCTATTCAAGAACAAATTGATCGTAAAGTCCACAGTAAACGACGACTACTTTG 327
Db 593 TACTAAAGGCTATTCAAGAACAAATTGATCGTAAAGTCCACAGTAAACGACGACTACTTTG 652
QY 328 AGGTCATTGATTTTGCAGGCGATGCAACCAATTACTGATCGAAACGGCAAGTCTACTTTG 387
Db 653 AGGTCATTGATTTTGCAGGCGATGCAACCAATTACTGATCGAAACGGCAAGTCTACTTTG 712
QY 388 CTGACAAAGATGTTTGGTTAACTTGCCTGACCAACCTGTCCAAGAAATTTTGTAAAGCG 447
Db 713 CTGACAAAGATGTTTGGTTAACTTGCCTGACCAACCTGTCCAAGAAATTTTGTAAAGCG 772
QY 448 GACATGTCGCGTTAGACCATATAAGAAAAAACAATACAAAACCAAGCAATCTGTTG 507
Db 773 GACATGTCGCGTTAGACCATATAAGAAAAAACAATACAAAACCAAGCAATCTGTTG 832
QY 508 ATGTGGAATATACTGTACAGTTTACTCCCTTAAACCTGTATGATGATTTTCAGACCAAGTC 567
Db 833 ATGTGGAATATACTGTACAGTTTACTCCCTTAAACCTGTATGATGATTTTCAGACCAAGTC 892
QY 568 TCAAAGATCTAAGCTATTGAAACACTAGCTATCGGTGACCACTCATCTCAAGAAT 627
Db 893 TCAAAGATCTAAGCTATTGAAACACTAGCTATCGGTGACCACTCATCTCAAGAAT 952
QY 628 TACTAGCTCAAGCACAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTATGAAC 687
Db 953 TACTAGCTCAAGCACAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTATGAAC 1012
QY 688 GTGACTCTCAATCGTCACTCATGCAATGACATTTTCGTAGATTTTACCAATGGATC 747
Db 1013 GTGACTCTCAATCGTCACTCATGCAATGACATTTTCGTAGATTTTACCAATGGATC 1072
QY 748 AAGAGTTTACTTACCGTGTAAAAATCGGAAACAAGCTTATAGCATCAATAAAAAATCTG 807
Db 1073 AAGAGTTTACTTACCGTGTAAAAATCGGAAACAAGCTTATAGCATCAATAAAAAATCTG 1132
QY 808 GTCTGAATGAAGAAATAAACAACACTGACTGATCTCTGAGAAATATTACGTCCTTAAAA 867
Db 1133 GTCTGAATGAAGAAATAAACAACACTGACTGATCTCTGAGAAATATTACGTCCTTAAAA 1192
QY 868 AAGGGAAAAAGCGTATGATCCCTTTGATCGGAGTCACTTGAACCTGTTTACCATCAAT 927

Db 1193 AAGGGGAAAGCGGTATGATCCCTTTGATCGCAGTCACCTTGAAACTGTTCCACATCAAAAT 1252
 QY 928 AGCTTGATGTCGATACCAACGAATTCCTAAAAAGTCAGCAGCTCTTAAACAGCTAGCGAAC 987
 Db 1253 AGCTTGATGTCGATACCAACGAATTCCTAAAAAGTCAGCAGCTCTTAAACAGCTAGCGAAC 1312
 QY 988 GTAACCTAGACTTCAGAGATTTATACGATCCTCGTGATAGGCTAAACTACTCTACAACA 1047
 Db 1313 GTAACCTAGACTTCAGAGATTTATACGATCCTCGTGATAGGCTAAACTACTCTACAACA 1372
 QY 1048 ATCTCGATGCTTTGCTATATGAGACTATACCTTAACCTGGAAGTAGAGATAATCAAG 1107
 Db 1373 ATCTCGATGCTTTGCTATATGAGACTATACCTTAACCTGGAAGTAGAGATAATCAAG 1432
 QY 1108 ATGACACCAACCGTATCATACCGTTTATATGCGCAGCGACCCGAGGAGAGAAATGCTA 1167
 Db 1433 ATGACACCAACCGTATCATACCGTTTATATGCGCAGCGACCCGAGGAGAGAAATGCTA 1492
 QY 1168 GCTACCATTTAGCTGGT 1187
 Db 1493 GCTATCATTTAGCTATGAT 1512

RESULT 9

AAx80492
 ID AAX80492 standard; cDNA; 1242 BP.
 AC AAX80492;
 XX 17-OCT-2003 (revised)
 DT 26-AUG-1999 (first entry)
 XX Streptococcus equisimilis native streptokinase encoding cDNA.
 XX Streptococcus; streptokinase; fibrin-dependent plasminogen activator;
 KW nSK; rSK; bacterial; blood clot; thrombotic condition;
 KW myocardial infarction; venous thrombosis; pulmonary embolism;
 KW cerebral thrombosis; graft thrombosis; arterial thrombosis; ss.
 XX Streptococcus dysgalactiae subsp. equisimilis.
 XX WO9931247-A1.
 FN 24-JUN-1999.
 PD 15-DEC-1998; 98WO-US026694.
 PF 15-DEC-1997; 97US-0069497P.
 PR (HARD) HARVARD COLLEGE.
 PA Reed GL;
 XX WPI; 1999-395183/33.
 DR P-PSDB; AAY24794.
 XX N-terminally deleted streptokinase.
 PT Claim 44; Page 58-60; 73pp; English.

The present invention describes an isolated bacterial protein that induces fibrin-dependent plasminogen activation in a pharmaceutical composition for dissolving blood clots. Also described are: (1) a composition comprising an isolated modified streptokinase, the modification being removal of amino acid residues in the amino terminus; (2) a method for dissolving a blood clot in a subject, comprising administering to the subject a fibrin-dependent streptokinase protein; a nucleic acid (1) encoding a modified bacterial streptokinase; (3) an expression vector comprising (1); and (4) a host cell transformed with the expression vector of (3). The pharmaceutical composition comprising a bacterial fibrin-dependent plasminogen activator is useful for dissolving blood clots in patients with a thrombotic condition, e.g. myocardial infarction, venous thrombosis, pulmonary embolism, cerebral thrombosis,

CC graft thrombosis and arterial thrombosis. The modified streptokinase can also be used in non-human mammals. Streptokinase activation of CC plasminogen is at least 10-fold, preferably 100-fold greater in the presence of fibrin than in the absence of fibrin. The modified streptokinase has at least one amino acid substitution that inactivates a CC substrate site for proteolytic cleavage. This reduces the rate of CC degradation of the streptokinase at least two-fold. The present sequence CC encodes native streptokinase (nSK). (Updated on 17-OCT-2003 to CC standardise OS field)

XX Sequence 1242 BP; 424 A; 267 C; 237 G; 314 T; 0 U; 0 Other;

Query Match 73.5%; Score 1132.6; DB 2; Length 1242;

Best Local Similarity 98.8%; Pred. No. 0;

Matches 1141; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 33 ATAGCTGGTCTCAATGGCTACTAGATCGTCTTGTGTAATAATACAGCCAAATGGTGT 92

Db 1 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAAATAGTTGT 60

QY 93 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTTAGTCTTAAATTTTGAATC 152

Db 61 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTTAGTCTTAAATTTTGAATC 120

QY 153 GATCTAACATCACGACCTGCTCATGGAGGAAAGACAGACGAGCTTTAAGTCCAAATCA 212

Db 121 GATCTAACATCACGACCTGCTCATGGAGGAAAGACAGACGAGCTTTAAGTCCAAATCA 180

QY 213 AAACCAATTTGCTACTGATAGTGGCGCATGTCTACATAAACTTGAGAAAGCTGACTTA 272

Db 181 AAACCAATTTGCTACTGATAGTGGCGCATGTCTACATAAACTTGAGAAAGCTGACTTA 240

QY 273 AAGGCTATTCAAGAACAAATTTGATCGCTAACGTCACAGTAACGACGACTTTGAGGTC 332

Db 241 AAGGCTATTCAAGAACAAATTTGATCGCTAACGTCACAGTAACGACGACTTTGAGGTC 300

QY 333 ATTGATTTTCAAGAGCATGCAACCATTTACTGATGCGAAACGCGAAGTCTACTTTGCTG 392

Db 301 ATTGATTTTCAAGAGCATGCAACCATTTACTGATGCGAAACGCGAAGTCTACTTTGCTG 360

QY 393 AAAGATGGTTCGGTAAACCTTTGCCGACCCCAACCTGTGCCAAGAAATTTTGTCTAAGCGG 452

Db 361 AAAGATGGTTCGGTAAACCTTTGCCGACCCCAACCTGTGCCAAGAAATTTTGTCTAAGCGG 420

QY 453 GTGCGGTTAGACCATATAAAGAAAAACCAATACAAAACCAAGGAAATCTGTGTGATG 512

Db 421 GTGCGGTTAGACCATATAAAGAAAAACCAATACAAAACCAAGGAAATCTGTGTGATG 480

QY 513 GAATATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCAGGCTCAAA 572

Db 481 GAATATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCAGGCTCAAA 540

QY 573 GATACTAAGCTATTGAAAAACATAGCTATCGGTGACACCATCATCTCAAGAAATTTACTA 632

Db 541 GATACTAAGCTATTGAAAAACATAGCTATCGGTGACACCATCATCTCAAGAAATTTACTA 600

QY 633 GCTCAAGCAAAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTATGAACGTCAC 692

Db 601 GCTCAAGCAAAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTATGAACGTCAC 660

QY 693 TCCTCAATCGTCACTCATGACAAATGACATTTTCGGTACGATTTTACCAATGGATCAAGAG 752

Db 661 TCCTCAATCGTCACTCATGACAAATGACATTTTCGGTACGATTTTACCAATGGATCAAGAG 720

QY 753 TTTTACTTACCGTCTTAAAAAATCGGAAACAGCTTTATAGGATCAATAAAAAATCTGGTCTG 812

Db 721 TTTTACTTACCGTCTTAAAAAATCGGAAACAGCTTTATAGGATCAATAAAAAATCTGGTCTG 780

QY 813 AATGAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTAACGTCCTTAAAAAAGGG 872

Db 781 AATGAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTAACGTCCTTAAAAAAGGG 840

QY 873 GAAAGCCGTATGATCCCTTTGATGCGACGTCATCTTGAACCTGTTCCACCATCAATAGGTT 932

Db 841 GAAAGCCGATGATCCCTTTGATCGAGTCACCTTGAACCTGTTCAACATCAATAGCTT 900
 QY 933 GATGTCGATACCAACGAATTTGCTAAAGAGTGAGCAGCTCTTAAACAGCTAGCGAAGTAAC 992
 Db 901 GATGTCGATACCAACGAATTTGCTAAAGAGTGAGCAGCTCTTAAACAGCTAGCGAAGTAAC 960
 QY 993 TTAGACTTCAGAGATTTATACGATCTCTGATGAAGGCTAAACTACTCTACAACAATCTC 1052
 Db 961 TTAGACTTCAGAGATTTATACGATCTCTGATGAAGGCTAAACTACTCTACAACAATCTC 1020
 QY 1053 GATGCTTTGGTATTATGACTATACCTTAACTGGAAGATAGAGGATAATCAGATGAC 1112
 Db 1021 GATGCTTTGGTATTATGACTATACCTTAACTGGAAGATAGAGGATAATCAGATGAC 1080
 QY 1113 ACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAAGGAGAGAGTCTAGCTAC 1172
 Db 1081 ACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAAGGAGAGAGTCTAGCTAT 1140
 QY 1173 CATTTAGCTGGTGGT 1187
 Db 1141 CATTTAGCTATGAT 1155

RESULT 10

AAA37633
 ID AAA37633 standard; DNA; 1245 BP.

AC AAA37633;

DT 15-SEP-2003 (revised)
 DT 13-OCT-2000 (first entry)

XX S. equisimilis streptokinase coding sequence.

KW Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
 KW plasminogen; human; fibrinectin; thrombolytic therapy;
 KW cardiovascular disorder; fibrinectin; ss.

OS Streptococcus dysgalactiae subsp. equisimilis.

FH Key Location/Qualifiers
 FT 1. .1245
 CDS /*tag= a
 FT /product= "streptokinase"

EP1024192-A2.

XX 02-AUG-2000.

XX 23-DEC-1999; 99EP-00310541.

XX 24-DEC-1998; 98IN-DE003825.

XX (COUL) CSIR: COUNCIL SCI IND RES.

PI Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
 PI Yadav M;

XX WPI; 2000-516032/47.

DR P-PSDB; AAY90282.

XX Hybrid streptokinase-fibrin binding domain polypeptides useful for
 PT thrombolytic therapy comprises a streptokinase fused with fibrin binding
 PT domains of human fibrinectin.

XX Example 3; Fig 3; 58pp; English.

CC This sequence represents the human Streptococcus equisimilis
 CC streptokinase coding sequence. The invention relates to a hybrid
 CC plasminogen activator (PA) comprises a polypeptide fusion between
 CC streptokinase (SK), which are capable of plasminogen (PG) activation, and
 CC fibrin binding regions of human fibrinectin, which are from fibrin

CC binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA possesses the
 CC ability to bind with fibrin independently and also characteristically
 CC retains a PG activation ability which becomes evident only after a
 CC pronounced duration, or lag, after exposure of the PA to a suitable
 CC animal or human PG. The hybrid streptokinase-fibrin binding domain
 CC polypeptides are useful in thrombolytic therapy for various kinds of
 CC cardiovascular disorders. The hybrids have enhanced fibrin selectivity as
 CC well as kinetics of plasminogen activation that are distinct from that of
 CC natural streptokinase in being characterised by a temporary delay, or lag
 CC of several minutes in the natural rate of the catalytic conversion of
 CC plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins
 CC can bind tightly with fibrin in blood clots soon after introduction into
 CC the vascular system without significantly activating the circulating
 CC blood plasminogen to plasmin, thus aiding in the localisation of the
 CC plasminogen activation process to the site of pathological thrombus. This
 CC overcomes systemic plasminogen activation encountered during clinical use
 CC of streptokinase. (Updated on 15-SEP-2003 to standardise OS field)
 XX

SQ Sequence 1245 BP; 426 A; 267 C; 237 G; 315 T; 0 U; 0 Other;

Query Match 73.5%; Score 1132.6; DB 3; Length 1245;

Best Local Similarity 98.8%; Pred. No. 0;

Matches 1141; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 33 ATAGCTGTCCTGAATGGCTACTAGATCGTCTCTGTAAATAACAGCAATTTGTTGTT 92

Db 1 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCCAATTAGTTGTT 60

QY 93 AGCGTTGCTGTACTGTTGAGGGGACGAATCAAGACATTAAGCTTTAAATTTTGAATC 152

Db 61 AGCGTTGCTGTACTGTTGAGGGGACGAATCAAGACATTAAGCTTTAAATTTTGAATC 120

QY 153 GATCTAACNTCAGCACCTGCTCATGGAGGAGGACAGAGCAAGCTTAAGTCCAAATCA 212

Db 121 GATCTAACNTCAGCACCTGCTCATGGAGGAGGACAGAGCAAGCTTAAGTCCAAATCA 180

QY 213 AAACATTTTGTCTACTGATAGTGGCGCATGTACATAAACTTGAGAAGCTGACTTACTA 272

Db 181 AAACATTTTGTCTACTGATAGTGGCGCATGTACATAAACTTGAGAAGCTGACTTACTA 240

QY 273 AAGGCTATTCAAGAACCAATTTGATCGCTAAACGTTCCACAGTAAACGACGACTACTTTGAGTGC 332

Db 241 AAGGCTATTCAAGAACCAATTTGATCGCTAAACGTTCCACAGTAAACGACGACTACTTTGAGTGC 300

QY 333 ATTGATTTTGCACGCGATGCAACCATTTACTGATCGAAACGGCAAGGCTCTACTTTGCTGAC 392

Db 301 ATTGATTTTGCACGCGATGCAACCATTTACTGATCGAAACGGCAAGGCTCTACTTTGCTGAC 360

QY 393 AAGATGTTTCGGTAACCTTTGCCGACCCCACTCTGCCAAGAAATTTTGTAAAGCGGACAT 452

Db 361 AAGATGTTTCGGTAACCTTTGCCGACCCCACTCTGCCAAGAAATTTTGTAAAGCGGACAT 420

QY 453 GTGCGCGTTAGACCATATATAAGAAAAACCAATACAAAAACCAAGCGAAATCTGTTGATGTG 512

Db 421 GTGCGCGTTAGACCATATATAAGAAAAACCAATACAAAAACCAAGCGAAATCTGTTGATGTG 480

QY 513 GAATATATCTGTACAGTTTACTTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTCAAA 572

Db 481 GAATATATCTGTACAGTTTACTTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTCAAA 540

QY 573 GATACTAAGCTATTGAAAAACACTAGCTTATCGGTGACACCATCATCTCAAGAAATTTACTA 632

Db 541 GATACTAAGCTATTGAAAAACACTAGCTTATCGGTGACACCATCATCTCAAGAAATTTACTA 600

QY 633 GCTCAAGCACAAGCAATTTTAAACAAAAACCAAGGCTATACGATTTTATGAACGTCGAC 692

Db 601 GCTCAAGCACAAGCAATTTTAAACAAAAACCAAGGCTATACGATTTTATGAACGTCGAC 660

QY 693 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 752

Db 661 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 720

QY 753 TTTTACTACCGTGTATAAAAAATCGGGAACAAGCTTTATAGGATCAATATAAAAAATCTGGTCTG 812

Db 721 TTTACTTACCGTGTAAATAATCGGAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 780
 Qy 813 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAGGG 872
 Db 781 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAGGG 840
 Qy 873 GAAAGCCGATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCACCATCAAAATAGCTT 932
 Db 841 GAAAGCCGATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCACCATCAAAATAGCTT 900
 Qy 933 GATGTCGATACCAACGAATTTGCTTAAAGTGACGCTCTTAAACAGCTAGCAACGTAAC 992
 Db 901 GATGTCGATACCAACGAATTTGCTTAAAGTGACGCTCTTAAACAGCTAGCAACGTAAC 960
 Qy 993 TTAGACTTTCAGAGATTTATACGATCTCTGATGAGCTTAAAGCTTAACTCTACACAAATCTC 1052
 Db 961 TTAGACTTTCAGAGATTTATACGATCTCTGATGAGCTTAAAGCTTAACTCTACACAAATCTC 1020
 Qy 1053 GATGCTTTTGGTATTATGACATATACCTTAACTGGAAAAAGTAGAGGATATACAGATGAC 1112
 Db 1021 GATGCTTTTGGTATTATGACATATACCTTAACTGGAAAAAGTAGAGGATATACAGATGAC 1080
 Qy 1113 ACCAACCGTATCATACCGTTTATATGCGCAAGCGACCCGAAAGGAGAGATGCTAGCTAC 1172
 Db 1081 ACCAACCGTATCATACCGTTTATATGCGCAAGCGACCCGAAAGGAGAGATGCTAGCTAT 1140
 Qy 1173 CATTTAGCTGGTGGT 1187
 Db 1141 CATTTAGCTATGAT 1155

RESULT 11

ABA05546
 ID ABA05546 standard; cDNA; 1254 BP.
 XX
 AC ABA05546;
 XX
 DT 26-FEB-2002 (first entry)
 XX
 DE Streptokinase cDNA.
 XX
 KW Streptokinase; cerebroprotective; cardiant; gene therapy; fusion;
 KW vasodilator; thrombolytic; angina; myocardial infarction; stroke;
 KW gene therapy; maxadilan; ss.
 XX
 OS Unidentified.
 XX
 PN WO200185100-A2.
 XX
 PD 15-NOV-2001.
 XX
 PF 10-MAY-2001; 2001WO-US015209.
 XX
 PR 11-MAY-2000; 2000US-00569920.
 XX
 PA (GENO) GEN HOSPITAL CORP.
 XX
 PI Reddy VB, Lerner E;
 XX
 DR WPI; 2002-062184/08.
 XX
 PT New fusion protein or conjugate, useful for treating unstable angina,
 PT acute myocardial infarction or stroke, comprises a vasodilator
 PT polypeptide and a thrombolytic polypeptide, or active fragments of the
 PT polypeptides.
 XX
 PS Example 1; Fig 2; 37pp; English.
 XX
 CC The invention relates to a fusion protein or a conjugate comprising a
 CC vasodilator polypeptide, or its active fragment, and a thrombolytic
 CC polypeptide or its active fragment. The protein is useful for treating a
 CC subject suffering from a partially or totally occluded blood vessel,

CC causing unstable angina, acute myocardial infarction or stroke. The
 CC polynucleotide encoding the polypeptide is useful in gene therapy. The
 CC vasodilator action of the protein allows for the use of lower doses of a
 CC thrombolytic while maintaining the clot dissolving effectiveness of the
 CC thrombolytic, and the use of lower doses of the thrombolytic reduces the
 CC associated side effects. The present sequence is the streptokinase cDNA
 CC used in the construction of a Maxadilan-Streptokinase fusion protein.
 CC Maxadilan is a vasodilator peptide produced by the salivary gland of the
 CC New World sand fly
 XX
 SQ Sequence 1254 BP; 426 A; 271 C; 241 G; 316 T; 0 U; 0 Other;

Query Match 73.5%; Score 1132.6; DB 6; Length 1254;
 Best Local Similarity 98.8%; Pred. No. 0;
 Matches 1141; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 33 ATAGCTGGTCTGTAATGGCTACTAGATCGTCCTTCTGTAATAACAGCAATTTGTTGTTT 92
 Db 7 ATTGCTGGACCTGAGTGGCTGCTAGACCGCTCATCTGTCAACAACAGCAATTTGTTGTT 66
 Qy 93 AGCGTTCTGCTACTGTTGAGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 152
 Db 67 AGCGTTCTGCTACTGTTGAGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 126
 Qy 153 GATCTAACATCACGACCTGCTCATGAGAGAAAGACAGAGCAAGGCTTAAGTCCAAATCA 212
 Db 127 GATCTAACATCACGACCTGCTCATGAGAGAAAGACAGAGCAAGGCTTAAGTCCAAATCA 186
 Qy 213 AAACCATTTGCTACTGATAGTGGCGGATGTCATATAAAGCTTGAGAAAGCTGACTACTA 272
 Db 187 AAACCATTTGCTACTGATAGTGGCGGATGTCATATAAAGCTTGAGAAAGCTGACTACTA 246
 Qy 273 AAGCTATTCAAGAAACAAATTTGATCGCTTAAGCTTCAAGTACGACGACTCTTTGAGTGC 332
 Db 247 AAGCTATTCAAGAAACAAATTTGATCGCTTAAGCTTCAAGTACGACGACTCTTTGAGTGC 306
 Qy 333 ATTGATTTTCAAGCGATGCAACCATTTACTGATCGAAACCGCAAGGCTTACTTTGCTGAC 392
 Db 307 ATTGATTTTCAAGCGATGCAACCATTTACTGATCGAAACCGCAAGGCTTACTTTGCTGAC 366
 Qy 393 AAAGATGGTTCGGTAACTTTGCCGACCAACCTGTCCAAAGAAATTTTGTCTAAGCGGACAT 452
 Db 367 AAAGATGGTTCGGTAACTTTGCCGACCAACCTGTCCAAAGAAATTTTGTCTAAGCGGACAT 426
 Qy 453 GTGCGCGTTAGACCATATAAGAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 512
 Db 427 GTGCGCGTTAGACCATATAAGAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 486
 Qy 513 GAATATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTCAAA 572
 Db 487 GAATATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTCAAA 546
 Qy 573 GATACTAAGCTATTGAAACAACCTAGCTATCGGTGACACCAATCACTCAAGAAATTTACTA 632
 Db 547 GATACTAAGCTATTGAAACAACCTAGCTATCGGTGACACCAATCACTCAAGAAATTTACTA 606
 Qy 633 GCTCAAGCAACAAGCATTTTAAACAAAACCAAGCGCTATACGATTTATGAACGTCAC 692
 Db 607 GCTCAAGCAACAAGCATTTTAAACAAAACCAAGCGCTATACGATTTATGAACGTCAC 666
 Qy 693 TCCTCAATCGTCACTCATGACAAATGACATTTTCCTGCTAGATTTTACCAATGGATCAAGAG 752
 Db 667 TCCTCAATCGTCACTCATGACAAATGACATTTTTCCTGCTAGATTTTACCAATGGATCAAGAG 726
 Qy 753 TTTACTTACCGTGTAAAAAATCGGGAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 812
 Db 727 TTTACTTACCGTGTAAAAAATCGGGAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 786
 Qy 813 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATAATTACGTCCTTAAAAAGGG 872
 Db 787 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATAATTACGTCCTTAAAAAGGG 846
 Qy 873 GAAAGCCGATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCACCATCAAAATAGCTT 932

Db 847 GAAAGCCGATGATCCCTTTGATCGAGTCACCTTGAACCTGTTCCACATCAAAATACGTT 906
 Qy 933 GATGTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTTAACAGCTAGCGAACGTAAC 992
 Db 907 GATGTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTTAACAGCTAGCGAACGTAAC 966
 Qy 993 TTAGACTTCAGAGATTTATACGATCCCTCGTGATAGGCTAAACTACTCTACAACAATCTC 1052
 Db 967 TTAGACTTCAGAGATTTATACGATCCCTCGTGATAGGCTAAACTACTCTACAACAATCTC 1026
 Qy 1053 GATGCTTTTGGTATTATGGACTATACCTTAACCTGGAAGCTAGAGGATATATCAGATGAC 1112
 Db 1027 GATGCTTTTGGTATTATGGACTATACCTTAACCTGGAAGCTAGAGGATATATCAGATGAC 1086
 Qy 1113 ACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAAAGGAGAGATGCTAGCTAC 1172
 Db 1087 ACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAAAGGAGAGATGCTAGCTAC 1146
 Qy 1173 CATTTAGCTGGTGGT 1187
 Db 1147 CATTTAGCTATGAT 1161

RESULT 12

ID ABA05547
 AC ABA05547 standard; DNA; 8893 BP.

XX ABA05547;

DT 26-FEB-2002 (first entry)

XX Maxadilan-streptokinase fusion protein plasmid pTVB3maxstk.

DE Streptokinase; cerebroprotective; cardiant; gene therapy; fusion;
 KW vasodilator; thrombolytic; angina; myocardial infarction; stroke;
 KW gene therapy; maxadilan; sand fly; plasmid; ds.

XX Lutzomyia longipalpis.

OS Unidentified.

OS Synthetic.

OS Chimeric.

XX WO200185100-A2.

PN 15-NOV-2001.

PD 10-MAY-2001; 2001WO-US015209.

XX 11-MAY-2000; 2000US-00569920.

XX (GENO) GEN HOSPITAL CORP.

PA Reddy VB, Lerner E;

XX WPI; 2002-062184/08.

DR New fusion protein or conjugate, useful for treating unstable angina,

XX acute myocardial infarction or stroke, comprises a vasodilator

PT polypeptide and a thrombolytic polypeptide, or active fragments of the

PT polypeptides.

XX Example 1; Fig 3; 37pp; English.

PS The invention relates to a fusion protein or a conjugate comprising a

XX vasodilator polypeptide, or its active fragment, and a thrombolytic

CC polypeptide or its active fragment. The protein is useful for treating a

CC subject suffering from a partially or totally occluded blood vessel,

CC causing unstable angina, acute myocardial infarction or stroke. The

CC polynucleotide encoding the polypeptide is useful in gene therapy. The

CC polynucleotide action of the protein allows for the use of lower doses of a

CC thrombolytic while maintaining the clot dissolving effectiveness of the

CC thrombolytic, and the use of lower doses of the thrombolytic reduces

CC associated side effects. The present sequence is the plasmid used for the
 CC preparation of Maxadilan-Streptokinase fusion protein. It contains
 CC maxadilan cDNA, which encodes a vasodilator peptide produced by the
 CC salivary gland of the New World sand fly, and streptokinase cDNA
 XX

SQ Sequence 8893 BP; 2366 A; 2106 C; 2246 G; 2175 T; 0 U; 0 Other;

Query Match 73.5%; Score 1132.6; DB 6; Length 8893;
 Best Local Similarity 98.8%; Pred. No. 0;
 Matches 1141; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 33 ATAGCTGGTCTCGAATGGCTACTAGATCGTCTTCTGTAAATAACAGCAATGGTTGTT 92
 Db 5935 ATTCTCGACCTCGAGTGGCTCTAGACCGTCCATCTGTCAACAACAGCAATAGTTGTT 5994
 Qy 93 AGCGTCTCGTACTGTGTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 152
 Db 5995 AGCGTCTCGTACTGTGTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 6054
 Qy 153 GATCTAACATCAACGACCTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAGTCCAAATCA 212
 Db 6055 GATCTAACATCAACGACCTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAGTCCAAATCA 6114
 Qy 213 AAACCATTTGCTACTGATAGTGGCGCGATGTCACATAAACTTGAGAAAGCTGACTACTA 272
 Db 6115 AAACCATTTGCTACTGATAGTGGCGCGATGTCACATAAACTTGAGAAAGCTGACTACTA 6174
 Qy 273 AAGGCTATTCAAGAAACAATGATCGCTTAACGTCACAGTACAGGAGCTACTTTGAGGTC 332
 Db 6175 AAGGCTATTCAAGAAACAATGATCGCTTAACGTCACAGTACAGGAGCTACTTTGAGGTC 6234
 Qy 333 ATTGATTTTGCAGCGGATGCAACCATTTACTCGATCGAAACGCGAGGCTACTTTGCTGAC 392
 Db 6235 ATTGATTTTGCAGCGGATGCAACCATTTACTCGATCGAAACGCGAGGCTACTTTGCTGAC 6294
 Qy 393 AAAGATGGTTTGGTAACTTTGCCGACCCCAACCTCTGCAAGAAATTTTGTCAAGCGGACAT 452
 Db 6295 AAAGATGGTTTGGTAACTTTGCCGACCCCAACCTCTGCAAGAAATTTTGTCAAGCGGACAT 6354
 Qy 453 GTGCGGCTTAGACCATATAAAGAAACCAATAAACAAGCGAAATCTGTTGATGTG 512
 Db 6355 GTGCGGCTTAGACCATATAAAGAAACCAATAAACAAGCGAAATCTGTTGATGTG 6414
 Qy 513 GAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAAGCGGCTCTCAA 572
 Db 6415 GAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAAGCGGCTCTCAA 6474
 Qy 573 GATACTAAGCTATTGAAACCACTAGCTATCGGTGACACCATCATCTCTCAAGAAATTAATA 632
 Db 6475 GATACTAAGCTATTGAAACCACTAGCTATCGGTGACACCATCATCTCTCAAGAAATTAATA 6534
 Qy 633 GCTCAAGCAAAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTTATGAACGTGAC 692
 Db 6535 GCTCAAGCAAAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTTATGAACGTGAC 6594
 Qy 693 TCCTCAATCGTCACTCATGACAAATGACATTTTCGGTAGCATTTTACCAATGGATCAAGAG 752
 Db 6595 TCCTCAATCGTCACTCATGACAAATGACATTTTCGGTAGCATTTTACCAATGGATCAAGAG 6654
 Qy 753 TTTTACTTTACCGTGTATAAATTCGGGAACAAGCTTATAGGATCAATAAATAATCTGGTCTG 812
 Db 6655 TTTTACTTTACCGTGTATAAATTCGGGAACAAGCTTATAGGATCAATAAATAATCTGGTCTG 6714
 Qy 813 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAAGTCTCTTAAAAAAGGG 872
 Db 6715 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAAGTCTCTTAAAAAAGGG 6774
 Qy 873 GAAAGCCGATGATGATCCCTTTGATCGCAGTCACTTGAACCTGTTTCAACATCAATAGCTT 932
 Db 6775 GAAAGCCGATGATGATCCCTTTGATCGCAGTCACTTGAACCTGTTTCAACATCAATAGCTT 6834
 Qy 933 GATGTCGATACCAACGAATTTGCTTAAAGTGAGCAGCTCTTAAACAGCTAGGACGTAAC 992

Db 6835 GATGTCGATACCAACGAATTGCTTAAAAAGTGAGCAGCTCTTAACAGCTAGCAACGTAAC 6894
 QY 993 TTAGACTTCAGAGATTATACGATCCTCGTGATAAGGCTAAACTACTCTACACAATCTC 1052
 Db 6895 TTAGACTTCAGAGATTATACGATCCTCGTGATAAGGCTAAACTACTCTACACAATCTC 6954
 QY 1053 GATGCTTTGGTATTATGACTATATACCTTAACTGGAAAAAGTAGAGGATATACAGATGAC 1112
 Db 6955 GATGCTTTGGTATTATGACTATATACCTTAACTGGAAAAAGTAGAGGATATACAGATGAC 7014
 QY 1113 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAAGAGAGAGATGCTAGCTAC 1172
 Db 7015 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAAGAGAGAGATGCTAGCTAT 7074
 QY 1173 CATTTAGCTGGTGGT 1187
 Db 7075 CATTTAGCTATGAT 7089

RESULT 13
 AAX16632
 ID AAX16632 standard; DNA; 1242 BP.
 AC AAX16632;
 DT 17-OCT-2003 (revised)
 DT 04-MAY-1999 (first entry)
 XX Streptococcus equisimilis native streptokinase encoding DNA.
 XX Streptococcus equisimilis H46A; streptokinase; mutant; fibrinolytic;
 KW plasma clot; hydrolysis; haemolytic Streptococcus; plasmin;
 KW serine protease; fibrin; blood clot; thrombolytic;
 KW vascular thromboembolytic symptom; acute myocardial infarction;
 KW fibrinolysis; resistance; ds.
 XX Streptococcus dysgalactiae subsp. equisimilis.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1242
 FT /tag= a
 FT /transl_except= (pos:40..42,aa:Asn)
 FT /note= "no stop codon given"
 XX
 FN US5876999-A.
 XX
 XX 02-MAR-1999.
 XX
 XX 06-DEC-1995; 95US-00568393.
 XX
 XX 06-DEC-1995; 95US-00568393.
 XX
 XX (NASC-) NAT SCI COUNCIL.
 XX
 XX Wu H;
 XX
 XX WPI; 1999-189643/16.
 DR P-PSDB; AAW94664.
 XX
 XX Mutant streptokinase polypeptide - useful as plasmin-resistant
 PT thrombolytic agent.
 XX
 XX Claim 1; Col 7-10; 17pp; English.
 XX
 XX The present invention describes a mutant streptokinase (SK) polypeptide
 CC in which at least one of the amino acids in the Pro58-Lys59-Ser60-Lys61
 CC segment of the corresponding native SK is replaced by another amino acid.
 CC The present sequence encodes native SK. SK is a secretory protein of
 CC haemolytic Streptococcus able to activate human plasminogen (HPlg) to
 CC plasmin (HPlm), which is a serine protease able to catalyse the
 CC hydrolysis of fibrin in blood clots. The SK is useful as a thrombolytic
 CC agent in the treatment of vascular thromboembolytic symptoms such as
 CC acute myocardial infarction. Compared with wild-type SK, the K59E mutant

CC is more resistant to degradation by human plasmin and is more effective
 CC both in acting as a fibrolytic agent and in activating human plasminogen.
 CC (Updated on 17-OCT-2003 to standardise OS field)
 XX
 SQ Sequence 1242 BP; 423 A; 268 C; 237 G; 314 T; 0 U; 0 Other;
 Query Match 73.4%; Score 1131; DB 2; Length 1242;
 Best Local Similarity 98.7%; Pred. No. 0;
 Matches 1140; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
 QY 33 ATAGCTGGTCTCAATGGCTACTAGATCGTCTTCTGTGTAAATAACAGCCAAATGGTTGTT 92
 Db 1 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCCACCAACGCAATAGTTGTT 60
 QY 93 AGCGTTTCTGGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTTTGAATC 152
 Db 61 AGCGTTTCTGGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTTTGAATC 120
 QY 153 GATCTAACATCACGACCTGCTCATGAGGAAAGACAGACGAAGCTTTAAGTCCAAAATCA 212
 Db 121 GATCTAACATCACGACCTGCTCATGAGGAAAGACAGACGAAGCTTTAAGTCCAAAATCA 180
 QY 213 AAACCAATTTGCTACTGATAGTGGCGGATGTCAATAAACTTGAGAAAGCTGACTTACTA 272
 Db 181 AAACCAATTTGCTACTGATAGTGGCGGATGTCAATAAACTTGAGAAAGCTGACTTACTA 240
 QY 273 AAGGCTATTCAAGAACAAATTTGATCGCTTAACGCTCCACAGTAACGCACTACTTTGAGGTC 332
 Db 241 AAGGCTATTCAAGAACAAATTTGATCGCTTAACGCTCCACAGTAACGCACTACTTTGAGGTC 300
 QY 333 ATTGATTTTCAAGAGCATGCAACCATTTACTGATGCGAAACCGCAAGTCTACTTTGCTGAC 392
 Db 301 ATTGATTTTCAAGAGCATGCAACCATTTACTGATGCGAAACCGCAAGTCTACTTTGCTGAC 360
 QY 393 AAAGATGGTTCGGTAACTTTGCGGACCCCAACCTGTCCAAGAAATTTTGTAAAGCGGACAT 452
 Db 361 AAAGATGGTTCGGTAACTTTGCGGACCCCAACCTGTCCAAGAAATTTTGTAAAGCGGACAT 420
 QY 453 GTGCGGTTAGACCATATATAAGAAAAACCAATACAAAACCAAGGAAATCTGTTGATGTG 512
 Db 421 GTGCGGTTAGACCATATATAAGAAAAACCAATACAAAACCAAGGAAATCTGTTGATGTG 480
 QY 513 GAATATACGTACAGTTTACTCCCTTAAACCCCTGATGACCATTTTCAGACCAGGCTCTCAA 572
 Db 481 GAATATACGTACAGTTTACTCCCTTAAACCCCTGATGACCATTTTCAGACCAGGCTCTCAA 540
 QY 573 GATACTAAGCTATTGAAAAACATAGCTATCGGTGACACCATCATCATCTCAAGAAATTA 632
 Db 541 GATACTAAGCTATTGAAAAACATAGCTATCGGTGACACCATCATCATCTCAAGAAATTA 600
 QY 633 GCTCAAGCAAAAGCATTTTAAACCAAAACCAAGGCTTATACGATTTATGAACGTTGAC 692
 Db 601 GCTCAAGCAAAAGCATTTTAAACCAAAACCAAGGCTTATACGATTTATGAACGTTGAC 660
 QY 693 TCCTCAATCGTCACTCATGACATGATTTTCGGTACGATTTTACCAATGGATCAAGAG 752
 Db 661 TCCTCAATCGTCACTCATGACATGATTTTCGGTACGATTTTACCAATGGATCAAGAG 720
 QY 753 TTTTACTTACCGTGTAAAAAATCGGAAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 812
 Db 721 TTTTACTTACCGTGTAAAAAATCGGAAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 780
 QY 813 AATGAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTATCGTCTTAAAAAAGGG 872
 Db 781 AATGAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTATACGCTCTTAAAAAAGGG 840
 QY 873 GAAAACCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCCACATCAATAGCTT 932
 Db 841 GAAAACCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCCACATCAATAGCTT 900
 QY 933 GATGTGATACCAACCAATTTGCTTAAAGAGGAGCAGCTCTTAAACAGCTAGCAGCACTAAC 992
 Db 901 GATGTGATACCAACCAATTTGCTTAAAGAGGAGCAGCTCTTAAACAGCTAGCAGCACTAAC 960

Qy	993	TTAGACTTCAGAGATTTATACGATCCTCGTGATTAAGGCTAAACTCTCTACAAACAATCTC	1051
Db	961	TTAGACTTCAGAGATTTATACGATCCTCGTGATTAAGGCTAAACTCTCTACAAACAATCTC	1020
Qy	1053	GATCGTTTTGGTATATGGACTATACCTTAACCTGGAAGTAGAGGATATATCAGATGAC	1112
Db	1021	GATCGTTTTGGTATATGGACTATACCTTAACCTGGAAGTAGAGGATATATCAGATGAC	1080
Qy	1113	ACCAACCGTATCATAAACGGTTTTATATGGCAAGCGACCCGAAGGAGAGAGAATGCTAGCTAC	1172
Db	1081	ACCAACCGTATCATAAACGGTTTTATATGGCAAGCGACCCGAAGGAGAGAGAATGCTAGCTAT	1140
Qy	1173	CATTAGCTGGTGGT	1187
Db	1141	CATTAGCCTATGAT	1155
RESULT 14			
AA	AX16633	ID	AX16633 standard; DNA; 1242 BP.
XX	XX	AC	AX16633;
XX	XX	04-MAY-1999	(first entry)
XX	XX	Streptococcus equisimilis	mutant streptokinase K59E encoding DNA.
XX	XX	Streptococcus equisimilis H46A;	streptokinase; mutant; fibrinolytic; plasma clot; hydrolysis; haemolytic Streptococcus; plasminogen; plasmin; serine protease; fibrin; blood clot; thrombolytic;
XX	XX	vascular thromboembolytic symptom;	acute myocardial infarction; fibrinolysis; resistance; ds.
OS	OS	Streptococcus dysgalactiae subsp. equisimilis.	Synthetic.
XX	XX	Key	Location/Qualifiers
PH	FT	CDS	1..1242
FT	FT		/tag= a
FT	FT		/trans_except= (pos:40..42,aa:Asn)
FT	FT		/note= "no stop codon given"
XX	XX	US5876999-A.	
XX	XX	02-MAR-1999.	
XX	XX	06-DEC-1995;	95US-00568393.
XX	XX	06-DEC-1995;	95US-00568393.
XX	XX	(NASC-)	NAT SCI COUNCIL.
XX	XX	Wu H;	
XX	XX	WPI;	1999-189643/16.
XX	XX	P-PSDB;	AAK94665.
XX	XX	Mutant streptokinase polypeptide -	useful as plasmin-resistant thrombolytic agent.
XX	XX	Claim 4;	Col 11-14; 17pp; English.
XX	XX	The present invention describes a mutant streptokinase (SK)	polypeptide in which at least one of the amino acids in the Pro58-Lys59-Ser60-Lys61 segment of the corresponding native SK is replaced by another amino acid.
XX	XX	The present sequence encodes mutant SK K59E.	SK is a secretory protein of haemolytic Streptococcus able to activate human plasminogen (Hplg) to plasmin (Hplm), which is a serine protease able to catalyse the hydrolysis of fibrin in blood clots.
XX	XX	The SK is useful as a thrombolytic agent in the treatment of vascular thromboembolytic symptoms such as acute myocardial infarction.	Compared with wild-type SK, the K59E mutant is more resistant to degradation by human plasmin and is more effective

CC	both in acting as a fibrolytic agent and in activating human plasminogen
XX	
SQ	Sequence 1242 BP; 421 A; 268 C; 239 G; 314 T; 0 U; 0 Other;
	Query Match 73.2%; Score 1127.8; DB 2; Length 1242;
	Best Local Similarity 98.5%; Pred. No. 0;
	Matches 1138; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
QY	33 ATAGCTCGTCTCGAATGGCTACTAGATCGTCTCTTCTGTAATAACAGCCAAATTGGTTGTT 92
DB	1 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCACCAACAGCCAAATTAGTTGTT 60
QY	93 AGCGTTCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGGAAATC 152
DB	61 AGCGTTCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGGAAATC 120
QY	153 GATCTAACATCAGACACCTGCTCATGGAGGAAAGACAGACGAAGGCCTTAAGTCCAAATCA 212
DB	121 GATCTAACATCAGACACCTGCTCATGGAGGAAAGACAGACGAAGGCCTTAAGTCCGAATCA 180
QY	213 AAACCAATTTGCTACTGATAGTGGCGCGATGTCACATAAACTTGAGAAAGCTGACTTACTA 272
DB	181 AAACCAATTTGCTACTGATAGTGGCGCGATGTCACATAAACTTGAGAAAGCTGACTTACTA 240
QY	273 AAGCTATTCAAGAACAAATTTGATCGCTAACAGTCAACAGTAAACGACGACTACTTTGAGGTC 332
DB	241 AAGCTATTCAAGAACAAATTTGATCGCTAACAGTCAACAGTAAACGACGACTACTTTGAGGTC 300
QY	333 ATTGATTTTCGACAGCGATGCAACCATTTACTGATCGAAACGGCAAGGTCTACTTTGCTGAC 392
DB	301 ATTGATTTTCGACAGCGATGCAACCATTTACTGATCGAAACGGCAAGGTCTACTTTGCTGAC 360
QY	393 AAAGATGGTTCCGGTAACTCTTCCGACCCCAACCTGTCCAAGAAATTTTGTCTAAAGCGACAT 452
DB	361 AAAGATGGTTCCGGTAACTCTTCCGACCCCAACCTGTCCAAGAAATTTTGTCTAAAGCGACAT 420
QY	453 GTGGCGGTTAGACCATATAAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 512
DB	421 GTGGCGGTTAGACCATATAAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 480
QY	513 GAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTACAGACAGGTCTCAAA 572
DB	481 GAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTACAGACAGGTCTCAAA 540
QY	573 GATACTAAGCTATTGAAAACACTTAGCTATCGGTGACCATCACTCTCAAGAATTTACTTA 632
DB	541 GATACTAAGCTATTGAAAACACTTAGCTATCGGTGACCATCACTCTCAAGAATTTACTTA 600
QY	633 GCTCAAGACACAAAGCAATTTTAAACAAAAACCCAGGCTATACGATTTTATGAACGTGAC 692
DB	601 GCTCAAGACACAAAGCAATTTTAAACAAAAACCCAGGCTATACGATTTTATGAACGTGAC 660
QY	693 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTTACGATTTTATCAATGGATCAAGAG 752
DB	661 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTTACGATTTTATCAATGGATCAAGAG 720
QY	753 TTTTACTTACCGTCTTAAATAACGGGACAGCTTATAGGATCAATATAAATACTGGTCTG 812
DB	721 TTTTACTTACCGTCTTAAATAACGGGACAGCTTATAGGATCAATATAAATACTGGTCTG 780
QY	813 AATGAAGAAATAAACCAACACTGACCTGATCTCTGAGAAATATATACGCTCTTAAAAAAGG 872
DB	781 AATGAAGAAATAAACCAACACTGACCTGATCTCTGAGAAATATATACGCTCTTAAAAAAGG 840
QY	873 GAAAGCCGTTATGATCCCTTTGATCCGAGTCACTTGAAACTGTFTTCAACATCAAAATACGTT 932
DB	841 GAAAGCCGTTATGATCCCTTTGATCCGAGTCACTTGAAACTGTFTTCAACATCAAAATACGTT 900
QY	933 GATGTCGATACCAACGAATTTGCTTAAAGGTGACGAGCTCTTAAACAGTACGGAACGTAC 992
DB	901 GATGTCGATACCAACGAATTTGCTTAAAGGTGACGAGCTCTTAAACAGTACGGAACGTAC 960
QY	993 TTAGACTTCAGAGATTTTATACGATCTCTGTTGTAAGGCTTAAACTACTCTTACCAACTCTC 1052

Db	961		TTAGACTTCAGAGATTTATACGATCCCTCGTGATAAGGCTAAACTACTCTACAAATCTC	1020
Qy	1053		GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAAGTAGAGGATATCAGATGAC	1112
Db	1021		GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAAGTAGAGGATAATCAGATGAC	1080
Qy	1113		ACCAACCGTATCATAAACGGTTTATATGGCAAGCGACCCGAGAGAGAATGCTAGCTAC	1172
Db	1081		ACCAACCGTATCATAAACGGTTTATATGGCAAGCGACCCGAGAGAGAATGCTAGCTAT	1140
Qy	1173		CATTTAGCTGGTGGT	1187
Db	1141		CATTTAGCTATGAT	1155

RESULT 15	
ADM01294	
ID	ADM01294 standard; DNA; 7057 BP.
XX	
AC	ADM01294;
XX	
DT	01-JUL-2004 (first entry)
XX	
XX	Plasmid pVTPA-Estrep DNA sequence.
DE	
XX	
KW	vector; plastid; artificial intergene region; plant;
KW	transplastomic angiosperm; agronomic property; stress resistance;
KW	rbcl gene; ds.
XX	
OS	Unidentified.

FN	XX	W02004029256-A2.
XX		
PD	XX	08-APR-2004.
XX		
PF	XX	15-SEP-2003; 2003WO-CU000009.
XX		
PR	XX	27-SEP-2002; 2002CU-00000208.
XX		
PA	XX	(INGG-) CENT ING GENETICA & BIOTECNOLOGIA.
XX		
PI	XX	Selman-Housein Sosa G, Aguilar Cabeza E, Gonzalez Quintero ADC;
PI	XX	Ramos Gonzalez O;
XX		
DR	XX	WPI; 2004-316131/29.
XX		
PT	XX	DNA vector for transformation and expression in plastids, useful e.g. for
PT	XX	producing pharmaceutical proteins or improving agronomic properties, has
PT	XX	gene inserted in artificial intergene region.
XX		
PS	XX	Example 10; SEQ ID NO 26; 74bp; Spanish.

The invention relates to a DNA vector (A) for stable transformation and expression of genes (I) in plants, where (I) is inserted in an artificial intergene region (AIR) formed by combining two 5'-untranslated regions (5'-UTRs) of genes that transcribe in different directions and are derived from plants of different divisions or classes. (A) are used to produce transplastomic angiosperms that have improved agronomic properties (e.g. resistance to biotic or abiotic stress) or express a very wide range of agricultural, veterinary, pharmaceutical, nutritional or industrial products, e.g. enzymes, vaccinating antigens, cytokines or immunoglobulins. Use of (A) eliminates the need for a transposon for gene insertion; inserted genes do not require promoters and terminators; and the structure of flanking sequences in (A) ensures universal applicability. Also any selection marker in (A) can be eliminated by homologous recombination. (A) provides efficient and stable expression of genes without causing any functional alterations. This sequence represents the plasmid pVTPA-Estrep, an example of the vector of the invention.

Query Match	72.7%	Score 1120.2	DB 12	Length 7057
Best Local Similarity	97.6%	Pred. No. 2.2e-312		
Matches 1137; Conservative	0;	Mismatches 28;	Indels	0; Gaps
Qy	23	ATATACCATGATAGCTGGTCCGTGAATGGCTACTAGATCGTCCTTCTGCTTAATACAGCCA	82	
Db	2013	ATAAACCATGATTGCTGGACCTGAGTGCGCTGCTAGACCCTGCATCTCTCAACAACAGCCA	2072	
Qy	83	ATTGGTTGTTAGCGTTCCTGCTACTGTTGAGGGGACGAATCAAGACATATTAGTCTTAAATT	142	
Db	2073	ATTAGTTGTTAGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATATTAGTCTTAAATT	2132	
Qy	143	TTTTGAAATCGATCTAATCATCAGACTGCTCATGAGAGGAAAGACAGCAAGCTTTAAG	202	
Db	2133	TTTTGAAATTTGACCTTAACATCAGACCCTGCTCATGAGAGGAAAGACAGCAAGCTTTAAG	2192	
Qy	203	TCCAAAATCAAACCATTTTCTACTGATGATGCGCGATGTCACATFAAATTGAGAAAGC	262	
Db	2193	TCCAAAATCAAACCATTTTCTACTGATGATGCGCGATGCCACATFAAATTGAAAAAGC	2252	
Qy	263	TGACTTACTAAAGGCTATTTCAGAACAATTTGATCGCTTAAACGTCACACAGTAACGACGACTA	322	
Db	2253	TGACTTACTAAAGGCTATTTCAGNAACAATTTGATCGCTTAAACGTCACAGTAACGACGACTA	2312	
Qy	323	CTTTGAGGTCATTGATTTTTCAGACGATCAACCATTTACTGATCGAAACCGCAAGGTCCTA	382	
Db	2313	CTTTGAGGTCATTGATTTTTCAGACGATCAACCATTTACTGATCGAAACCGCAAGGTCCTA	2372	
Qy	383	CTTTGCTGACAAAGATGGTTTCGGTAACTTTGCCGACCCCAACCTGTCCAGAAATTTTGCT	442	
Db	2373	CTTTGCTGACAAAGATGGTTTCGGTAACTTTGCCGACCCCAACCTGTCCAGAAATTTTGCT	2432	
Qy	443	AAGCGGACATGTGCGCTTAGACCATATAAGAAAAACCAATACAAAAACCAAGCGAAATC	502	
Db	2433	AAGCGGACATGTGCGCGTTAGACCATATAAGAAAAACCAATACAAAAACCAAGCGAAATC	2492	
Qy	503	TGTTGATGTGGAATATATCTGTACAGTTTACTTCCCTTAAACCTCGATGACGATTTTCAGACC	562	
Db	2493	TGTTGATGTGGAATATATCTGTACAGTTTACTTCCCTTAAACCTCGATGACGATTTTCAGACC	2552	
Qy	563	AGGCTCTCAAAGATACTTAAGCTATTGAAAAACATAGCTATCGGTGACACCATCATCTCA	622	
Db	2553	AGGCTCTCAAAGATACTTAAGCTATTGAAAAACATAGCTATCGGTGACACCATCATCTCA	2612	
Qy	623	AGAAATCTAGCTCAAGCACAAAGCATTTTAAACAAAAACCAACCCAGCGCTATACGATTTA	682	
Db	2613	AGAAATCTAGCTCAAGCACAAAGCATTTTAAACAAAAACCAACCCAGCGCTATACGATTTA	2672	
Qy	683	TGAACGTGACTCCTCAATCGTCACTCATGACAAATGACATTTTCOGTACGATTTTACCAAT	742	
Db	2673	TGAACGTGACTCCTCAATCGTCACTCATGACAAATGACATTTTCOGTACGATTTTACCAAT	2732	
Qy	743	GGATCAAGATTTTACTTTACCGTGTAAAAATCGGGAACAAGCTTTATAGGATCAATAAAAA	802	
Db	2733	GGATCAAGATTTTACTTTACCATGTGCAAAAAATCGGGAACAAGCTTTATAGGATCAATAAAAA	2792	
Qy	803	ATCTGCTCTCAATGAAGAAATAACACACTGACTGATCTCTGAGAAAAATATTACGTCCT	862	
Db	2793	ATCTGCTCTCAATGAAGAAATAACACACTGACTGATCTCTGAGAAAAATATTACGTCCT	2852	
Qy	863	TAAAAAGGGGAAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAAACTGTTCCACAT	922	
Db	2853	TAAAAAGGGGAAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAAACTGTTCCACAT	2912	
Qy	923	CAAAATAGTTGATGTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGCTAG	982	
Db	2913	CAAAATAGTTGATGTCAAACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGCTAG	2972	
Qy	983	CGAACGTAACCTTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTA	1042	
Db	2973	CGAACGTAACCTTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTA	3032	
Qy	1043	CAACAATCTCGATGCTTTTGGTATTTATGGACTATACCTTAACTGGAAAAAGTAGAGGATAA	1102	

Db	3033	CAACAATCTCGATGCTTTTGGTATTATGGACTATACCTTAACCTGGAAGTAGAGGATAA	3092
Qy	1103	TCACGATGACACCAACCGTATCATACCGTTTATATGGCAAGCGCCCGAAGGAGAGAA	1162
Db	3093	TCACGATGACACCAACCGTATCATACCGTTTATATGGCAAGCGCCCGAAGGAGAGAA	3152
Qy	1163	TGCTAGCTACCATTTAGCTGGTGGT	1187
Db	3153	TGCTAGCTATCATTTAGCCTATGAT	3177

Search completed: January 31, 2006, 18:17:49
Job time : 934.565 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2006, 16:24:19 ; Search time 272.069 Seconds
(without alignments)
10068.109 Million cell updates/sec

Title: US-09-940-235-9
Perfect score: 1541
Sequence: 1 ttgtttaacttaagaagg.....ccttcacgatgttgtag 1541

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 889780828 residues
Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_NA:*

1: /cgn2_6/ptodata/1/ina/1_COMB.seq:*

2: /cgn2_6/ptodata/1/ina/5_COMB.seq:*

3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*

4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*

5: /cgn2_6/ptodata/1/ina/H_COMB.seq:*

6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*

7: /cgn2_6/ptodata/1/ina/PP_COMB.seq:*

8: /cgn2_6/ptodata/1/ina/RE_COMB.seq:*

9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1134	73.6	2385	3	US-09-211-542A-1
2	1132.6	73.5	1242	2	US-09-211-542A-5
3	1131	73.4	1242	2	US-08-568-393B-1
4	1127.8	73.2	1242	2	US-08-568-393B-2
5	1118.6	72.6	2566	2	US-08-488-940-19
6	1111.8	72.1	1245	2	US-07-703-778D-1
7	1111.6	72.1	1257	2	US-07-854-596B-25
8	1110.2	72.0	1512	2	US-07-854-596B-27
9	1108.6	71.9	1317	2	US-07-854-596B-18
10	1108.6	71.9	1335	2	US-07-854-596B-18
11	1108.6	71.9	1458	2	US-07-854-596B-42
12	1108.6	71.9	1467	2	US-07-854-596B-46
13	1108.6	71.9	2589	2	US-07-854-596B-34
14	1084.4	70.4	1209	3	US-09-374-038-11
15	1084.4	70.4	1209	3	US-09-658-179-11
16	1084.4	70.4	1245	3	US-09-374-038-12
17	1084.4	70.4	1245	3	US-09-658-179-12
18	1083.8	70.3	1122	3	US-09-374-038-13
19	1083.8	70.3	1122	3	US-09-658-179-13
20	1083.8	70.3	1158	3	US-09-374-038-14
21	1083.8	70.3	1158	3	US-09-658-179-14
22	1078.8	70.0	2253	2	US-07-854-596B-39
23	1077.6	69.9	1119	2	US-07-854-596B-30
24	973	63.1	1068	3	US-09-211-542A-11

25	970	62.9	2208	3	US-09-211-542A-3	Sequence 3, Appli
26	761.2	49.4	1262	9	5240845-3	Patent No. 5240845
27	756.2	49.1	1242	9	5240845-2	Patent No. 5240845
28	450	29.2	450	3	US-09-211-542A-13	Sequence 13, Appl
29	330	21.4	7679	3	US-09-220-132-38	Sequence 38, Appl
30	330	21.4	7680	3	US-09-023-655-1289	Sequence 1289, Ap
31	330	21.4	7680	6	PCT-US95-09819-6	Sequence 6, Appli
32	330	21.4	7803	2	US-08-551-356-1	Sequence 1, Appli
33	330	21.4	7803	6	PCT-US93-13687-1	Sequence 1, Appli
34	330	21.4	8044	3	US-09-566-921-135	Sequence 135, App
35	326.8	21.2	7705	2	US-08-259-569-16	Sequence 16, Appl
36	326.8	21.2	7705	2	US-08-826-885-16	Sequence 16, Appl
37	326.8	21.2	7705	9	5455158-2	Patent No. 5455158
38	50.8	3.3	986	2	US-07-637-350A-8	Sequence 8, Appli
39	50.8	3.3	986	2	US-08-145-061-8	Sequence 8, Appli
C 40	49.8	3.2	7218	2	US-08-232-463-14	Sequence 14, Appl
C 41	45.4	2.9	55	9	5240845-38	Patent No. 5240845
42	45	2.9	45	9	5455158-17	Patent No. 5455158
43	45	2.9	45	9	5455158-18	Patent No. 5455158
44	43.4	2.8	53	9	5240845-33	Patent No. 5240845
45	41.8	2.7	53	9	5240845-34	Patent No. 5240845

ALIGNMENTS

RESULT 1
US-09-211-542A-1
; Sequence 1, Application US/09211542A
; Patent No. 6210667
; GENERAL INFORMATION:
; APPLICANT: Reed, Guy L.
; TITLE OF INVENTION: BACTERIAL FIBRIN-DEPENDENT PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROMBERG & SUNSTEIN, LLP
; STREET: 125 Summer Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 15-December-1998
; CLASSIFICATION: 1653
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/069,497
; FILING DATE: 15-December-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Attorney, Strimpel, Harriet M.
; REGISTRATION NUMBER: 37,008
; REFERENCE/DOCKET NUMBER: 1874/111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)443-9292
; TELEFAX: (617)443-0004
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2385 base pairs
; TYPE: nucleic acid
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2385
US-09-211-542A-1

Query Match 73.6%; Score 1134; DB 3; Length 2385;

Best Local Similarity 98.7%; Pred. No. 0; Matches 1143; Conservative 0; Mismatches 15; Indels 0; Gaps 0;			
Qy	30	ATGATAGCTCGGTCTCGAATGGCTACTAGATCGTCTTCTGTAAATAACAGCCAATTCGGTT	89
Db	1141	AGGATTGCTGGACCTGAGTGGCTCTAGACGGTCCATCTGTCAACAACGCCAATTAAGTT	1200
Qy	90	GTTAGCGTTGCTGGTACTGTTGAGGGGACGGAATCAAGACATTAAGTCTTAAATTTTTTGAA	149
Db	1201	GTTAGCGTTGCTGGTACTGTTGAGGGGACGGAATCAAGACATTAAGTCTTAAATTTTTTGAA	1260
Qy	150	ATCGATTAACATCAACGACCTGCTCATGAGGGAAGACAGACGAAGCTTTAAGTCCAAAA	209
Db	1261	ATCGATTAACATCAACGACCTGCTCATGAGGGAAGACAGACGAAGCTTTAAGTCCAAAA	1320
Qy	210	TCAAAACCAATTGCTACTGATAGTGGCGCGATGTCATAAACTTGAGAAGCTGACTTA	269
Db	1321	TCAAAACCAATTGCTACTGATAGTGGCGCGATGTCATAAACTTGAGAAGCTGACTTA	1380
Qy	270	CTAAAGGCTATTCAAGAACAAATTGATCGCTAAACGTCCACAGTACACGACGACTACTTTGAG	329
Db	1381	CTAAAGGCTATTCAAGAACAAATTGATCGCTAAACGTCCACAGTACACGACGACTACTTTGAG	1440
Qy	330	GTCAATTGATTTTGCAGCGAGTGCACCAATTACTGATCGAAACGGCAAGGCTCTACTTTGCT	389
Db	1441	GTCAATTGATTTTGCAGCGAGTGCACCAATTACTGATCGAAACGGCAAGGCTCTACTTTGCT	1500
Qy	390	GACAAAGATGGTTTCGGTAACCTTTCGCCACCACTCTGCCAAGAAATTTTTGCTAAAGCGGA	449
Db	1501	GACAAAGATGGTTTCGGTAACCTTTCGCCACCACTCTGCCAAGAAATTTTTGCTAAAGCGGA	1560
Qy	450	CATGTGCGGCTTAGACCATATAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGAT	509
Db	1561	CATGTGCGGCTTAGACCATATAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGAT	1620
Qy	510	GTGGAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGAAATTCAGACAGGTCCTC	569
Db	1621	GTGGAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGAAATTCAGACAGGTCCTC	1680
Qy	570	AAAGATACTTAAGCTATTGAAAAACACTAGCTATCGGTGACACCAATCAATCTCAAGAATTA	629
Db	1681	AAAGATACTTAAGCTATTGAAAAACACTAGCTATCGGTGACACCAATCAATCTCAAGAATTA	1740
Qy	630	CTAGCTCAAGCACAAGCATTTTAAACAAAAACACCCAGGCTATACGATTTATGAACGT	689
Db	1741	CTAGCTCAAGCACAAGCATTTTAAACAAAAACACCCAGGCTATACGATTTATGAACGT	1800
Qy	690	GACTCTCAATCGTCACTCANGACAATGACATTTTCGGTACGATTTTACCAATGGATCAA	749
Db	1801	GACTCTCAATCGTCACTCANGACAATTTTCGGTACGATTTTACCAATGGATCAA	1860
Qy	750	GAGTTTACTTACCGTGTAAAAATCGGGAACAGCTTATAGGATCAATAAANAATCTGGT	809
Db	1861	GAGTTTACTTACCGTGTAAAAATCGGGAACAGCTTATAGGATCAATAAANAATCTGGT	1920
Qy	810	CTGAATGAAGAAATAAACACACTGACCTCATCTCTGAGAAAAATTAACGTCCTTAAAAAA	869
Db	1921	CTGAATGAAGAAATAAACACACTGACCTCATCTCTGAGAAAAATTAACGTCCTTAAAAAA	1980
Qy	870	GGGGAAGCCGTATGATCCTTTGATCGAGTCACATTTGAAAATCTGTTCCACATCAAAATAC	929
Db	1981	GGGGAAGCCGTATGATCCTTTGATCGAGTCACATTTGAAAATCTGTTCCACATCAAAATAC	2040
Qy	930	GTTGATGTGATACCAAGAAATTTGCTAAAAAGTGAGCAGCTCTTTACAGCTGACGACGT	989
Db	2041	GTTGATGTGATACCAAGAAATTTGCTAAAAAGTGAGCAGCTCTTTACAGCTGACGACGT	2100
Qy	990	AACCTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAACTACTCTCAACAAT	1049
Db	2101	AACCTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAACTACTCTCAACAAT	2160
Qy	1050	CTCGATGCTTTTCGGTATTATGGACTATACCTTAACTGGAAGTAGGAGATAATACGAT	1109

```

Db      2161 CTCGATGCTTTTGGTATATATGGAGCTATACCTTTAACTCGGAAGAGTAGAGGATAATACGAT 2222
Qy      1110 GACACCAACCGTATCATCAACCGTTTATATGGGCAAGCGACCCGGAAGGAGAGAATGCTTACG 1169
Db      2221 GACACCAACCGTATCATCAACCGTTTATATGGGCAAGCGACCCGGAAGGAGAGAATGCTTACG 2280
Qy      1170 TACCATTTAGCTGGTGGT 1187
Db      2281 TATCATTTAGCCTATGAT 2298

RESULT 2
US-09-211-542A-5
; Sequence 5, Application US/09211542A
; Patent No. 6210667
; GENERAL INFORMATION:
; APPLICANT: Reed, Guy L.
; TITLE OF INVENTION: BACTERIAL FIBRIN-DEPENDENT PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROMBERG & SUNSTEIN, LLP
; STREET: 125 Summer Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/211,542A
; FILING DATE: 15-December-1998
; CLASSIFICATION: 1633
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/069,497
; FILING DATE: 15-December-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Attorney, Strimpel, Harriet M.
; REGISTRATION NUMBER: 37,008
; REFERENCE/DOCKET NUMBER: 1874/111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)443-9292
; TELEFAX: (617)443-0004
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1242 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1242
US-09-211-542A-5

```

	Query Match	73.5%	Score 1132.6	DB 3	Length 1242	
	Best Local Similarity	98.8%	Pred. No. 0			
	Matches 1141	Conservative 0	Mismatches 14	Indels 0	Gaps 0	
Qy	33	ATAGCTGGTCC	TGAATGGGCT	ACTAGATCGTCCT	TGTAAATACAGCCAA	TGGTTGTT 92
Db	1	ATTGCTGGAC	CTGAGTGGCT	GCTAGACCGTCC	ATCTGTCAACAGCCAA	TAGTTGTT 60
Qy	93	AGCGTGTCTG	GTACTGTTGAGGGG	AGCAATCAAGACA	TTAGTCTTAAATTTT	TGGAATC 152
Db	61	AGCGTGTCTG	TACTGTTGAGGGG	AGCAATCAAGACA	TTAGTCTTAAATTTT	TGGAATC 120
Qy	153	GATCTAACAT	CAAGACCTGCTCAT	TGGAGGAAACAGAG	CAAGGCTTTAGTCCAA	ATCA 212
Db	121	GATCTAACAT	CACAGACCTGCTCAT	TGGAGGAAACAGAG	CAAGGCTTTAGTCCAA	ATCA 180

```

QY 213 AAAACATTTGCTACTGATAGTGGCGGATGTCATATAAATTGAGAAAGCTGACTTACTA 272
Db 181 AAACCATTTGCTACTGATAGTGGCGGATGTCATATAAATTGAGAAAGCTGACTTACTA 240
QY 273 AAGGCTATTCAAGAAACAATTGATCGCTAAAGTCCACAGTAACGACGACTACTTTGAGGTC 332
Db 241 AAGGCTATTCAAGAAACAATTGATCGCTAAAGTCCACAGTAACGACGACTACTTTGAGGTC 300
QY 333 ATTGATTTTGGCAAGCGATGCAACATTTACTGATCGGAAACGGCAAGGCTCTACTTTGCTGAC 392
Db 301 ATTGATTTTGGCAAGCGATGCAACATTTACTGATCGGAAACGGCAAGGCTCTACTTTGCTGAC 360
QY 393 AAGATGTTTGGTACCTTTGCGGACCCAACTGTCCAAGAAATTTTGGCTAAGCGGACAT 452
Db 361 AAGATGTTTGGTACCTTTGCGGACCCAACTGTCCAAGAAATTTTGGCTAAGCGGACAT 420
QY 453 GTGCGGTTAGACCATATAAAGAAACCAATACCAACCAAGGAAATCTGTTGATGTG 512
Db 421 GTGCGGTTAGACCATATAAAGAAACCAATACCAACCAAGGAAATCTGTTGATGTG 480
QY 513 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTCAAA 572
Db 481 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTCAAA 540
QY 573 GATATTAAGCTATTGAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTA 632
Db 541 GATATTAAGCTATTGAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTA 600
QY 633 GCTCAAGCAAAAGCAATTTTAAACAAACACCCAGGCTATACGATTTATGAACGTGAC 692
Db 601 GCTCAAGCAAAAGCAATTTTAAACAAACACCCAGGCTATACGATTTATGAACGTGAC 660
QY 693 TCCTCAATCGTCACTCATGACATGACATTTTCGTTACGATTTTACCAATGGATCAAGAG 752
Db 661 TCCTCAATCGTCACTCATGACATGACATTTTCGTTACGATTTTACCAATGGATCAAGAG 720
QY 753 TTTACTTACCGTGTGTTAAATCGGAAACAGCTTTATAGGATCAATAAAAAATCTGGTCTG 812
Db 721 TTTACTTACCGTGTGTTAAATCGGAAACAGCTTTATAGGATCAATAAAAAATCTGGTCTG 780
QY 813 AATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAGGG 872
Db 781 AATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAGGG 840
QY 873 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTACCATCAAAATAGCTT 932
Db 841 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTACCATCAAAATAGCTT 900
QY 933 GATGTCGATACCAACGAATTTGCTTAAAGTGACGAGCTCTTAAACAGCTAGCGAACGTAAC 992
Db 901 GATGTCGATACCAACGAATTTGCTTAAAGTGACGAGCTCTTAAACAGCTAGCGAACGTAAC 960
QY 993 TTAGACTTCAGAGATTTATACGATCTCTGATGATAAGCTTAACTCTTACACAAATCTC 1052
Db 961 TTAGACTTCAGAGATTTATACGATCTCTGATGATAAGCTTAACTCTTACACAAATCTC 1020
QY 1053 GATGCTTTTGGTATTATGAGCTATATACCTTAACTGGAAGAGTAGAGGATATCAAGTAC 1112
Db 1021 GATGCTTTTGGTATTATGAGCTATATACCTTAACTGGAAGAGTAGAGGATATCAAGTAC 1080
QY 1113 ACCAACCGTATCATACCGTTTATATGCGGACCGGACCGGAGAGAGATGCTAGCTAC 1172
Db 1081 ACCAACCGTATCATACCGTTTATATGCGGACCGGACCGGAGAGAGATGCTAGCTAC 1140
QY 1173 CATTTAGCTGGTGGT 1187
Db 1141 CATTTAGCTGGTGGT 1155

```

RESULT 3

US-08-568-393B-1

; Sequence 1, Application US/08568393B

; Patent No. 5876999

```

;
; GENERAL INFORMATION:
; APPLICANT: Hua-Lin Wu
; APPLICANT: Guey-Yueh Shi
; TITLE OF INVENTION: Preparation of novel streptokinase
; TITLE OF INVENTION: mutants as improved thrombolytic agents
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jeing & Chang
; STREET: Two No. 5876999th Second Street, Suite 290
; CITY: San Jose
; STATE: California
; COUNTRY: USA
; ZIP: 95113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44MB
; MEDIUM TYPE: storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 6.1 on Window 3.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/568,393B
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Chi-Ping Chang
; REGISTRATION NUMBER: 37,798
; REFERENCE/DOCKET NUMBER:
; TELEPHONE: (408) 288-8585
; TELEFAX: (408) 288-8386
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1242 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE:
; HYPOTHEICAL: N
; ANTI-SENSE: N
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus equisimilis H46A
; INDIVIDUAL ISOLATE: Malke, H., Roe, B., and Ferretti, J. J.;
; INDIVIDUAL ISOLATE: "Nucleotide sequence of the streptokinase gene from Streptococcus equisimilis H46A" from Gene 34:357-362 (1985).
; CELL TYPE: Streptococcus equisimilis H46A
; US-08-568-393B-1

```

```

Query Match 73.4%; Score 1131; DB 2; Length 1242;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 1140; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 33 ATAGCTGGTCTCTGAATGGCTACTAGATCGTCTCTTCTGTAATAACACGCCAATTTGTTGTT 92
Db 1 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTCTCACCACACGCCAATTTGTTGTT 60
QY 93 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 152
Db 61 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 120
QY 153 GATCTAACATCAAGCTCTGCTCATGGAGGAAAGACAGACGCTTAAAGTCCAAATCA 212
Db 121 GATCTAACATCAAGCTCTGCTCATGGAGGAAAGACAGACGCTTAAAGTCCAAATCA 180
QY 213 AAACCATTTGCTACTGATAGTGGCGGATGTCATATAAATTGAGAAAGCTGACTTACTA 272
Db 181 AAACCATTTGCTACTGATAGTGGCGGATGTCATATAAATTGAGAAAGCTGACTTACTA 240
QY 273 AAGGCTATTCAAGAAACAATTGATCGCTTAAAGTCCACAGTAACGACGACTACTTTGAGGTC 332
Db 241 AAGGCTATTCAAGAAACAATTGATCGCTTAAAGTCCACAGTAACGACGACTACTTTGAGGTC 300
QY 333 ATTGATTTTGGCAAGCGATGCAACATTTACTGATCGGAAACGGCAAGGCTCTACTTTGCTGAC 392

```

```

Db 301 ATTGATTTTGAAGCGATGCAACCAATTAATGATCGAAACGGCAAGTCTACTTTGCTGAC 360
Qy 393 AAAGATGTTTGGTAACTTGGCCGACCCCACTCTGCAAGAAATTTTGTCTAAGCGGACAT 452
Db 361 AAAAGTGGTTGGTAACTTGGCCGACCCCACTCTGCAAGAAATTTTGTCTAAGCGGACAT 420
Qy 453 GTGCGCGTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTGTGATG 512
Db 421 GTGCGCGTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTGTGATG 480
Qy 513 GAATATAGTCTAGAGTTTACTCCCTTAAACCCCTGATGAGATTTGAGACAGTCTCAAA 572
Db 481 GAATATAGTCTAGAGTTTACTCCCTTAAACCCCTGATGAGATTTGAGACAGTCTCAAA 540
Qy 573 GATACTAGCTATTGAAACACCTAGCTATCGGTGACACCATCACATCTCAAGAAATTA 632
Db 541 GATACTAGCTATTGAAACACCTAGCTATCGGTGACACCATCACATCTCAAGAAATTA 600
Qy 633 GCTCAAGCACAAAGCATTTTAAACCAAAACCAACCCAGGCTATACGATTTTATGAACGTGAC 692
Db 601 GCTCAAGCACAAAGCATTTTAAACCAAAACCAACCCAGGCTATACGATTTTATGAACGTGAC 660
Qy 693 TCCTCAATCGTCACTATGACAAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 752
Db 661 TCCTCAATCGTCACTATGACAAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 720
Qy 753 TTTTACTTACCGTGTAAATAACCGGACCAAGCTTATAGGATCAATAAATAAATCTGCTCG 812
Db 721 TTTTACTTACCGTGTAAATAACCGGACCAAGCTTATAGGATCAATAAATAAATCTGCTCG 780
Qy 813 AATGAAGAAATAAACAACACCTGATCTCTGAGAAATATTAAGTCTCTTAAAAAAGG 872
Db 781 AATGAAGAAATAAACAACCTGATCTCTGAGAAATATTAAGTCTCTTAAAAAAGG 840
Qy 873 GAAAGCGGTATGATCCCTTTGATCGAGTCACTTGAACCTGTTCAACATCAAAATCGTT 932
Db 841 GAAAGCGGTATGATCCCTTTGATCGAGTCACTTGAACCTGTTCAACATCAAAATCGTT 900
Qy 933 GATGCGATCAACGAATGCTAAAGTAGAGCTCTTAAACAGCTAGCGAAGCTAAC 992
Db 901 GATGCGATCAACGAATGCTAAAGTAGAGCTCTTAAACAGCTAGCGAAGCTAAC 960
Qy 993 TTAGACTTCAGAGATTTATACGATCTCTGATAGAGCTTAAACTACTCTACAAATCTC 1052
Db 961 TTAGACTTCAGAGATTTATACGATCTCTGATAGAGCTTAAACTACTCTACAAATCTC 1020
Qy 1053 GATGCTTTTGGTATTATGAGCTATATACCTTAACTGGAAGAGTAGAGGATTAATCAGATGAC 1112
Db 1021 GATGCTTTTGGTATTATGAGCTATATACCTTAACTGGAAGAGTAGAGGATTAATCAGATGAC 1080
Qy 1113 ACCAAGCGTATCAACCGTTTATATGCGCAAGCGACCCGAAAGGAGAGAAATGCTAGCTAC 1172
Db 1081 ACCAAGCGTATCAACCGTTTATATGCGCAAGCGACCCGAAAGGAGAGAAATGCTAGCTAT 1140
Qy 1173 CATTTAGCTGGTGGT 1187
Db 1141 CATTTAGCTGGTGGT 1155

```

RESULT 4
US-08-568-393B-2

```

; Sequence 2, Application US/08568393B
; Patent No. 5876999
; GENERAL INFORMATION:
; APPLICANT: Hua-Lin Wu
; APPLICANT: Guey-Yueh Shi
; TITLE OF INVENTION: Preparation of novel streptokinase
; TITLE OF INVENTION: mutants as improved thrombolytic agents
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jeing & Chang
; STREET: Two No. 5876999th Second Street, Suite 290
; CITY: San Jose

```

```

; STATE: California
; COUNTRY: USA
; ZIP: 95113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44MB
; MEDIUM TYPE: storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 6.1 on Window 3.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/568,393B
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Chi-Ping Chang
; REGISTRATION NUMBER: 37,798
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (408) 288-8585
; TELEFAX: (408) 288-8386
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1242 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE:
; HYPOTHETICAL: N
; ANTI-SENSE: N
; FEATURE:
; NAME/KEY: SK-K59E
; LOCATION: DNA sequence No. 5876999174 and 175 have been changed
; LOCATION: from AA to GG, and PROTEIN sequence No. 587699959 has been changed
; LOCATION: from Lys to Glu.
; OTHER INFORMATION:
; US-08-568-393B-2

```

Query Match 73.2%; Score 1127.8; DB 2; Length 1242;

Best Local Similarity 98.5%; Pred. No. 8.7e-315; Indels 0; Gaps 0;
Matches 1138; Conservative 0; Mismatches 17;

```

Qy 33 ATAGCTGGTCTCGAATGGCTACTAGATCGTCTCTGTAAATAACACGCAATTTGTTGT 92
Db 1 ATTGCTGGACCTTGAGTGGCTGTAGACCGTCCATCTGTACCACACGCAATTTGTTGT 60
Qy 93 AGCGTTTGTGGTACTGTTGAGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 152
Db 61 AGCGTTTGTGGTACTGTTGAGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 120
Qy 153 GATCTAATCATCAGACCTGCTCATGAGGAAAGACAGACGAGCTTAAAGTCCAAATCA 212
Db 121 GATCTAATCATCAGACCTGCTCATGAGGAAAGACAGACGAGCTTAAAGTCCGGAATCA 180
Qy 213 AAACCAATTTGCTACTGATAGTGGCGGATGTCAATAAAGTTCAGAAAGCTGACTTACTA 272
Db 181 AAACCAATTTGCTACTGATAGTGGCGGATGTCAATAAAGTTCAGAAAGCTGACTTACTA 240
Qy 273 AAGGCTATTCAAGAACAAATTTGATCGCTAACAGTCCACAGTAACGACGCTACTTTAGGTC 332
Db 241 AAGGCTATTCAAGAACAAATTTGATCGCTAACAGTCCACAGTAACGACGCTACTTTAGGTC 300
Qy 333 ATTGATTTTGAAGCGATGCAACCATTAAGTATGATGCGAAACGGCAAGTCTACTTTGCTGAC 392
Db 301 ATTGATTTTGAAGCGATGCAACCATTAAGTATGATGCGAAACGGCAAGTCTACTTTGCTGAC 360
Qy 393 AAAGATGTTTCGTAAACCTTTGCCGACCCCAACCTGTCACAGAAATTTTGTCTAAGCGGACAT 452
Db 361 AAAGATGTTTCGTAAACCTTTGCCGACCCCAACCTGTCACAGAAATTTTGTCTAAGCGGACAT 420
Qy 453 GTGCGCGTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTGTGATG 512
Db 421 GTGCGCGTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTGTGATG 480

```

QY 513 GAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGGTCTCAAA 572
Db 481 GAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGGTCTCAAA 540
QY 573 GATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCTCAAGATTTACTA 632
Db 541 GATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCTCAAGATTTACTA 600
QY 633 GCTCAAGCACAAGCAATTTTAAACAAAACCCAGGCTATACGATTTTATGAACGTGAC 692
Db 601 GCTCAAGCACAAGCAATTTTAAACAAAACCCAGGCTATACGATTTTATGAACGTGAC 660
QY 693 TCCTCAATCGTCTCATGACCAATGATTTTCGTGACATTTTCAATGGATCAAGAG 752
Db 661 TCCTCAATCGTCTCATGACCAATGATTTTCGTGACATTTTCAATGGATCAAGAG 720
QY 753 TTTTACTTACCGTGTAAATTCGGACACGCTTATAGGATCAATAAAAAATCTGGTCTG 812
Db 721 TTTTACTTACCGTGTAAATTCGGACACGCTTATAGGATCAATAAAAAATCTGGTCTG 780
QY 813 AATGAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 872
Db 781 AATGAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 840
QY 873 GAAAGCCGTATGATCCCTTTGATCGCAGTCACCTTGAACTGTTTCAACATCAAAATACGTT 932
Db 841 GAAAGCCGTATGATCCCTTTGATCGCAGTCACCTTGAACTGTTTCAACATCAAAATACGTT 900
QY 933 GATGTCGATACCAACGAATGCTTAAAGTGAAGAGTCACTTTAAGCTAGCGAACGTAAC 992
Db 901 GATGTCGATACCAACGAATGCTTAAAGTGAAGAGTCACTTTAAGCTAGCGAACGTAAC 960
QY 993 TTAGACTTCAGAGATTTATACGATCTCTGATGATAAGGCTAAATCTCTTACACAAATCTC 1052
Db 961 TTAGACTTCAGAGATTTATACGATCTCTGATGATAAGGCTAAATCTCTTACACAAATCTC 1020
QY 1053 GATGCTTTTGGTATTATGGAATATACCTTAACTGGAAGTAGAGGATAATCAGATGAC 1112
Db 1021 GATGCTTTTGGTATTATGGAATATACCTTAACTGGAAGTAGAGGATAATCAGATGAC 1080
QY 1113 ACCAACCGTATCATACCGTTTATATGGAAGTAGAGGATAATCAGATGAC 1172
Db 1081 ACCAACCGTATCATACCGTTTATATGGAAGTAGAGGATAATCAGATGAC 1140
QY 1173 CATTTAGCTGGTGGT 1187
Db 1141 CATTTAGCTGGTGGT 1155

RESULT 5

US-08-488-940-19
; Sequence 19, Application US/08488940
; Patent No. 5854049
; GENERAL INFORMATION:
; APPLICANT: Reed, Guy L.
; TITLE OF INVENTION: PLASMIN-RESISTANT STREPTOKINASE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,940
; FILING DATE: 09-JUN-1995
; CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 05433/009001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2566 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-488-940-19

Query Match 72.6%; Score 1118.6; DB 2; Length 2566;
Best Local Similarity 98.6%; Pred. No. 6.2e-312;
Matches 1139; Conservative 0; Mismatches 14; Indels 2; Gaps 1;
QY 33 ATAGCTGGTCTGTAATGGCTACTAGATCGTCTCTGTGTAATAAATACAGCCAAATGGTGGT 92
Db 897 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAAACAGCCAAATTTAGTGGT 956
QY 93 AGCGTTGCTGTACTGTTTGGGGGACGAATCAAGACATTTAGTCTTAAATTTTTCGAATC 152
Db 957 AGCGTTGCTGTACTGTTTGGGGGACGAATCAAGACATTTAGTCTTAAATTTTTCGAATC 1016
QY 153 GATCTAACATCAACACCTGCTCATGAGAGGAAAGACAGACGAGGCTTAAGTCCAAATCA 212
Db 1017 GATCTAACATCAACACCTGCTCAT--AGGAAGACAGACGAGGCTTAAGTCCAAATCA 1074
QY 213 AAACCATTTGCTACTGATAGTGGCGGATGTCACATAAACTTGAGAAAGCTGACTACTA 272
Db 1075 AAACCATTTGCTACTGATAGTGGCGGATGTCACATAAACTTGAGAAAGCTGACTACTA 1134
QY 273 AAGGCTATTCAAGAACAAATGATCGCTTAAACGTCACAGTACGACGACTACTTTGAGGTC 332
Db 1135 AAGGCTATTCAAGAACAAATGATCGCTTAAACGTCACAGTACGACGACTACTTTGAGGTC 1194
QY 333 ATTGATTTTGCAGCGATGCAACCATTTACTGATCGAAACGGCAAGGCTCTACTTTGCTGAC 392
Db 1195 ATTGATTTTGCAGCGATGCAACCATTTACTGATCGAAACGGCAAGGCTCTACTTTGCTGAC 1254
QY 393 AAAGATGGTTCGGTAACTTTGCCGACCCCAACCTGTCCAAGAAATTTTGTCTAAGCGGACAT 452
Db 1255 AAAGATGGTTCGGTAACTTTGCCGACCCCAACCTGTCCAAGAAATTTTGTCTAAGCGGACAT 1314
QY 453 GTGGCGTTAGACCATATAAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 512
Db 1315 GTGGCGTTAGACCATATAAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 1374
QY 513 GAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGGTCTCAAA 572
Db 1375 GAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGGTCTCAAA 1434
QY 573 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTTACTA 632
Db 1435 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTTACTA 1494
QY 633 GCTCAAGCACAAGCAATTTTAAACAAAACCCAGGCTATACGATTTTATGAACGTGAC 692
Db 1495 GCTCAAGCACAAGCAATTTTAAACAAAACCCAGGCTATACGATTTTATGAACGTGAC 1554
QY 693 TCCTCAATCGTCACTCATGACCAATGACATTTTCGTAGGATTTTACCAATGGATCAAGAG 752
Db 1555 TCCTCAATCGTCACTCATGACCAATGACATTTTCGTAGGATTTTACCAATGGATCAAGAG 1614
QY 753 TTTTACTTACCGTGTAAATTCGGAAACAAAGCTTATAGGATCAATAAAAAATCTGGTCTG 812
Db 1615 TTTTACTTACCGTGTAAATTCGGAAACAAAGCTTATAGGATCAATAAAAAATCTGGTCTG 1674

Db 901 GATGTCACCAACGAATTGCTTAAAGCGAGCGCTCTTAAACAGCTAGCGAACGTAAC 960
Qy 993 TTAGACTTCAGAGATTATACGATCCTCGTGATAAGGCTAAAGCTTACTCTTACAACAATCTC 1052
Db 961 TTAGACTTCAGAGATTATACGATCCTCGTGATAAGGCTAAAGCTTACTCTTACAACAATCTC 1020
Qy 1053 GATGCTTTTGGTATTATGAGCTATACCTTTAACTGGAAAAGTAGAGGATAATCAGGATGAC 1112
Db 1021 GATGCTTTTGGTATTATGAGCTATACCTTTAACTGGAAAAGTAGAGGATAATCAGGATGAC 1080
Qy 1113 ACCAACCGTATCATACCGTTTATATGGCAACGAGCCCGAAGGAGAGAAATGCTAGCTAC 1172
Db 1081 ACCAACCGTATCATACCGTTTATATGGCAACGAGCCCGAAGGAGAGAAATGCTAGCTAT 1140
Qy 1173 CATTTAGCTGGTGGT 1187
Db 1141 CATTTAGCTATGAT 1155

RESULT 7
US-07-854-596B-25;
; Sequence 25, Application US/07854596B
; Patent No. 5434073
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith M
; APPLICANT: Hunter, Michael G
; APPLICANT: Czaplowski, Lloyd G
; TITLE OF INVENTION: Proteins and nucleic acids
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. John J. McDonnell
; STREET: Ten South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/854,596B
; FILING DATE: 03-JUN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,337
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1257 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..1257
; OTHER INFORMATION: /note= "Methionyl-streptokinase
; fusion protein"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 4..1248
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 4..1248
US-07-854-596B-25

Query Match 72.1%; Score 1111.6; DB 2; Length 1257;
Best Local Similarity 97.5%; Pred. No. 4.7e-310;
Matches 1129; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
Qy 30 ATGATAGCTGGTCTCGTAATGGCTACTAGATCGTCTTCTGTGTAATAACAGCAATTTGGTT 89
Db 4 ATGATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCAATTTAGTT 63
Qy 90 GTTAGCCGTTCTGGTACTGTTGAGGGGACGGAATCAAGACATTAGTCTTTAAATTTTTTGA 149
Db 64 GTTAGCCGTTCTGGTACTGTTGAGGGGACGGAATCAAGACATTAGTCTTTAAATTTTTTGA 123
Qy 150 ATCGATCTAAACATCAGCACCTGCTCATGAGAGAAAGACAGAGCAAGGCTTAAGTCCAAAA 209
Db 124 ATTGACCTTAACATCAGCACCTGCTCATGAGAGAAAGACAGAGCAAGGCTTAAGTCCAAAA 183
Qy 210 TCAAAACCATTTGCTACTGATAGTGGCGCGATGTCATATAAATTGAGAAAGCTGACTTA 269
Db 184 TCAAAACCATTTGCTACTGATAGTGGCGCGATGTCATATAAATTGAGAAAGCTGACTTA 243
Qy 270 CTAAAGGCTATTCAAGAAACAATTGATCGCTTAACGTCACAGTAAGGAGCTACTTTGAG 329
Db 244 CTAAAGGCTATTCAAGAAACAATTGATCGCTTAACGTCACAGTAAGGAGCTACTTTGAG 303
Qy 330 GTCATTGATTTTGCAGCGGATGCAACCATTTACTGATCGAAACGCAAGGCTCTACTTTGCT 389
Db 304 GTCATTGATTTTGCAGCGGATGCAACCATTTACTGATCGAAACGCAAGGCTCTACTTTGCT 363
Qy 390 GACAAAGATGGTTGGGTAACTTTGCCGACCCCAACTGTGCCAAGAAATTTTGTCTAAGCGGA 449
Db 364 GACAAAGATGGTTGGGTAACTTTGCCGACCCCAACTGTGCCAAGAAATTTTGTCTAAGCGGA 423
Qy 450 CATGTGCGCTTAGACCATATAAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGAT 509
Db 424 CATGTGCGCTTAGACCATATAAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGAT 483
Qy 510 GTGGAATATATCTGACAGTTTACTCCCTTAAACCTGATGACGATTTTACAGCCAGGCTC 569
Db 484 GTGGAATATATCTGACAGTTTACTCCCTTAAACCTGATGACGATTTTACAGCCAGGCTC 543
Qy 570 AAAGATATAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCTCAAGAATTA 629
Db 544 AAAGATATAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCTCAAGAATTA 603
Qy 630 CTAGCTCAAGCACAAGCAATTTTAAACAAAAACCCAGGCTATACGATTTATGAACGT 689
Db 604 CTAGCTCAAGCACAAGCAATTTTAAACAAAAACCCAGGCTATACGATTTATGAACGT 663
Qy 690 GACTCTCAATCGTCACTCATGACATGACATTTTCCGTAGGATTTTACCAATGGATCAA 749
Db 664 GACTCTCAATCGTCACTCATGACATGACATTTTCCGTAGGATTTTACCAATGGATCAA 723
Qy 750 GAGTTTACTTACCGTGTGTTAAAAATCGGGAACAAGCTTATAGGATCAATAAAAAATCTGGT 809
Db 724 GAGTTTACTTACCGTGTGTTAAAAATCGGGAACAAGCTTATAGGATCAATAAAAAATCTGGT 783
Qy 810 CTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAACGCTCTTAAAAAAA 869
Db 784 CTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAACGCTCTTAAAAAAA 843
Qy 870 GGGGAAAGCCGTATGATCCCTTTGATCGGAGTCACCTTGAACATGTTTCAACATCAATAC 929
Db 844 GGGGAAAGCCGTATGATCCCTTTGATCGGAGTCACCTTGAACATGTTTCAACATCAATAC 903
Qy 930 GTTGATGTCGATACCAACGAATTTGCTAAAAGTGAGCAGCTCTTAAACAGCTAGGACGT 989
Db 904 GTTGATGTCGATACCAACGAATTTGCTAAAAGTGAGCAGCTCTTAAACAGCTAGGACGT 963
Qy 990 AACTTAGCTTCAGAGATTTTATACGATCTCTCGTGATAAGGCTAAAGCTTACTCTTACAACAAT 1049
Db 964 AACTTAGCTTCAGAGATTTTATACGATCTCTCGTGATAGGCTAAAGCTTACTCTTACAACAAT 1023

254	ATATAAGAAATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTGCAACAACAGCCCAAT	313
Db		
85	TGGTTGTTTACGGCTTGCTGGTACTGTTTCAGGGGACGGAATCAAGACATTAGTCTCTTAAATTTT	144
Qy		
314	TAGTTGTTTAGCGTTGCTGGTACTGTTTCAGGGGACGGAATCAAGACATTAGTCTCTTAAATTTT	373
Db		
145	TTGAAATCGATCTAAACATCAGACTGCTCATGGAGGAAAACAGACCAAGGCTTTAAGTC	204
Qy		
374	TTGAAATTGACCTTAACATCAGACTGCTCATGGAGGAAAACAGACCAAGGCTTTAAGTC	433
Db		
205	CAAAATCAAAACCACTTTGCTACTGATAGTGGCGCGATGTCACTAAATCTTGAGAAAGCTG	264
Qy		

00 434 C A T T C A T T T G C T A C T G A T A N I S O C C O C A T O C C A C A T A T A T T C A T T A T T A C T G 493

QY	265	ACTTACTAAAGGCTATTTC	CAAGACAAATTGATCGCTAACGTC	CCACAGGTAAACGACGACTACT	324
DB	494	ACTTACTAAAGGCTATTTC	CAAGACAAATTGATCGCTAA	CGTCCACAGTAAACGACGACTACT	553
QY	325	TTGAGTGCAATTGATTTT	GCAAGCGATGCAACCAATTACTGAT	TGCAAAACGGCAAGGTCCTACT	384
DB	554	TTGAGTGCAATTGATTTT	GCAAGCGATGCAACCAATTACTGAT	TGCAAAACGGCAAGGTCCTACT	613
QY	385	TTGCTGCAAGAGATGGTT	TCGGTAACTTCGCGACCCAAACCTGT	CCAGAATATTTTGTCTAA	444
DB	614	TTGCTGCAAGAGATGGTT	TCGGTAACTTCGCGACCCAAACCTGT	CCAGAATATTTTGTCTAA	673
QY	445	GCGGACATGTCGCGTT	TAGACCATATATAAGAAACCAATACAAAA	CCAAAGCGGAATCTG	504
DB	674	GCGGACATGTCGCGTT	TAGACCATATATAAGAAACCAATACAAAA	CCAAAGCGGAATCTG	733
QY	505	TTGATGTGGAAATACCT	GTACAGTTTACTCCCTTAAACCCCTGAT	GACGATTCAGACCAG	564
DB	734	TTGATGTGGAAATACCT	GTACAGTTTACTCCCTTAAACCCCTGAT	GACGATTCAGACCAG	793
QY	565	GTCTCAAAGATACCTAAG	CAATTTGAAAACACTAGCTATCGGTG	ACACCATCACATCTCAAG	624
DB	794	GTCTCAAAGATACCTAAG	CAATTTGAAAACACTAGCTATCGGTG	ACACCATCACATCTCAAG	853
QY	625	AATTACTAGCTCAAGCA	CAAGCAATTTTAAACAAAAACCAACCCAGGCT	TATACGATTTATG	684
DB	854	AATTACTAGCTCAAGCA	CAAGCAATTTTAAACAAAAACCAACCCAGGCT	TATACGATTTATG	913
QY	685	AACGTGACTCCTCAAT	CGTCACTCATGACAAATGACATTTTC	CGTACGATTTTACCAGTGG	744
DB	914	AACGTGACTCCTCAAT	CGTCACTCATGACAAATGACATTTTC	CGTACGATTTTACCAGTGG	973
QY	745	ATCAAGAGTTTACTTTAC	TCGTTTAAATAACCGGGAACCAAGCTTTA	TAGGATCAATAAAAAAT	804
DB	974	ATCAAGAGTTTACTTTAC	TCGTTTAAATAACCGGGAACCAAGCTTTA	TAGGATCAATAAAAAAT	1033
QY	805	CTGCTCTGAATGAAGAA	ATAAAACAACTGACCTGATCTCTGAGAAAT	TATATCGTCCCTTA	864
DB	1034	CTGCTCTGAATGAAGAA	ATAAAACAACTGACCTGATCTCTGAGAAAT	TATATCGTCCCTTA	1093
QY	865	AAAAAGGGGAAAAGCC	GTATGATCCCTTTGATCGAGTCACTTGAA	ACTGTTTCAACCATCA	924
DB	1094	AAAAAGGGGAAAAGCC	GTATGATCCCTTTGATCGAGTCACTTGAA	ACTGTTTCAACCATCA	1153
QY	925	AATACGTTGATGCGAT	ACCNAACGAATTTGCTAAAAAGT	GACGACTCTTAAACGCTAGCG	984
DB	1154	AATACGTTGATGCGAT	ACCNAACGAATTTGCTAAAAAGT	GACGACTCTTAAACGCTAGCG	1213
QY	985	AACGTAACTTAGACTTC	CAGAGATTTATACGATCCTCGT	GATAAGGCTTAAACTACTCTTACA	1044
DB	1214	AACGTAACTTAGACTTC	CAGAGATTTATACGATCCTCGT	GATAAGGCTTAAACTACTCTTACA	1273
QY	1045	ACAATCTCGATGCTTTT	GGTATTTATGGAATATACCTTTAACT	TGGAAAAAGTAGAGGATAATC	1104
DB	1274	ACAATCTCGATGCTTTT	GGTATTTATGGAATATACCTTTAACT	TGGAAAAAGTAGAGGATAATC	1333
QY	1105	ACGATGACCAACCGGT	ATCATACCGTTTATATGGGCAACCGCAC	CCGAGGAGAGGAGATG	1164
DB	1334	ACGATGACCAACCGGT	ATCATACCGTTTATATGGGCAACCGCAC	CCGAGGAGAGGAGATG	1393

QY 1165 CTAGCTACCAATTTAGCTGGT 1187
 Db 1394 CTAGCTATCATTTAGCCTATGAT 1416

RESULT 9

US-07-854-596B-18
 ; Sequence 18, Application US/07854596B
 ; Patent No. 5434073
 ; GENERAL INFORMATION:
 ; APPLICANT: Dawson, Keith M
 ; APPLICANT: Hunter, Michael G
 ; APPLICANT: Czaplowski, Lloyd G
 ; TITLE OF INVENTION: Proteins and nucleic acids
 ; NUMBER OF SEQUENCES: 73
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Dr. John J. McDonnell
 ; STREET: Ten South Wacker Drive, Suite 3000
 ; CITY: Chicago
 ; STATE: IL
 ; COUNTRY: USA
 ; ZIP: 60606
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/854,596B
 ; FILING DATE: 03-JUN-1992
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: McDonnell, John J
 ; REGISTRATION NUMBER: 26,949
 ; REFERENCE/DOCKET NUMBER: 92,337
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 312-715-1234
 ; TELEFAX: 312-715-1234
 ; TELEX: 910-221-5317
 ; INFORMATION FOR SEQ ID NO: 18:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1317 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: 1..1317
 ; OTHER INFORMATION: /note= "OmpA fused to mature
 ; OTHER INFORMATION: streptokinase gene"
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 4..1308
 ; FEATURE:
 ; NAME/KEY: mat_peptide
 ; LOCATION: 4..1308
 ; US-07-854-596B-18

Query Match 71.9%; Score 1108.6; DB 2; Length 1317;
 Best Local Similarity 97.5%; Pred. No. 3.5e-309;
 Matches 1126; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
 QY 33 ATAGCTGGTCTCGAATGGCTACTAGATCGTCTTCTGTAAATAACAGCCAAATTTGGTTGTT 92
 Db 67 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAAGCCCAATTTAGTTGTT 126
 QY 93 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTTAGTCTTAAATTTTGTGAATC 152
 Db 127 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTTAGTCTTAAATTTTGTGAATC 186
 QY 153 GATCTAACATCATCAGCACTCTCATGAGGAGGAGAGAGAGAGGCTTTAGTCCAAAATCA 212

Db	187	GACCTAAACATCAGCACTGCTCATGGAGGAAAGACAGAGCAGGCTTAAGTCCAAAATCA	246
QY	213	AAACCATTTTGTCTACTGATAGTGGCGCGATGTACATAACTTGTGAGAAAGCTGACTACTA	272
Db	247	AAACCATTTTGTCTACTGATAGTGGCGCGATGCCACATAAACTTGAAGAGCTGACTACTA	306
QY	273	AAGGCTATTCAAGAACAAATTTGATTCGCTAAACGTCACAGTAACGAGCACTACTTTGAGGTC	332
Db	307	AAGGCTATTCAAGAACAAATTTGATTCGCTAAACGTCACAGTAACGAGCACTACTTTGAGGTC	366
QY	333	ATTGATTTTGAAGCGATGCAACCAATTAATCTGATCGAAACGCGCAAGGTCTACTTTCTGAC	392
Db	367	ATTGATTTTGAAGCGATGCAACCAATTAATCTGATCGAAACGCGCAAGGTCTACTTTCTGAC	426
QY	393	AAAGATGGTTCGGTAAACCTTCCGCGACCAACCTGTCGAAGATTTTTCCTAGCGGACAT	452
Db	427	AAAGATGGTTCGGTAAACCTTCCGCGACCAACCTGTCGAAGATTTTTCCTAGCGGACAT	486
QY	453	GTGCGCGTTAGACCATATAAAGAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG	512
Db	487	GTGCGCGTTAGACCATATAAAGAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG	546
QY	513	GAATATATCTGTACAGTTTACTCTCCCTTAAACCTGTATGACGATTTTCAGACCGAGTCTCAA	572
Db	547	GAATATATCTGTACAGTTTACTCTCCCTTAAACCTGTATGACGATTTTCAGACCGAGTCTCAA	606
QY	573	GATACTAAGCTATTGAAAACACTAGCTATCGGTGACCACTACATCTCAAGAAATTA	632
Db	607	GATACTAAGCTATTGAAAACACTAGCTATCGGTGACCACTACATCTCAAGAAATTA	666
QY	633	GCTCAAGCACAAAGCATTTTAAACAAAACCAACCCAGGCTATACGATTTTATGAACGTCAC	692
Db	667	GCTCAAGCACAAAGCATTTTAAACAAAACCAACCCAGGCTATACGATTTTATGAACGTCAC	726
QY	693	TCCTCAATCGTCACTCATGACAAATGCAATTTTCCGTACGATTTTACCAATGGATCAAGAG	752
Db	727	TCCTCAATCGTCACTCATGACAAATGCAATTTTCCGTACGATTTTACCAATGGATCAAGAG	786
QY	753	TTTACTTACCGTGTAAATAATCGGGAAACAAAGCTTATAGGATCAATAAATAATCTGGTCTG	812
Db	787	TTTACTTACCGTGTAAATAATCGGGAAACAAAGCTTATAGGATCAATAAATAATCTGGTCTG	846
QY	813	AATGAAGAAATAAACACACACCTGATCTCTGAGAAATATTAGTCTTTAAAAAGGG	872
Db	847	AATGAAGAAATAAACACACACCTGATCTCTGAGAAATATTAGTCTTTAAAAAGGG	906
QY	873	GAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTTGAAACTGTTTACCATCAAAATACGTT	932
Db	907	GAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTTGAAACTGTTTACCATCAAAATACGTT	966
QY	933	GATGTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAAC	992
Db	967	GATGTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAAC	1026
QY	993	TTAGACTTCAGAGATTTATAGATTCCTCGTGAATAGGCTAAACTACTCTACAAATCTC	1052
Db	1027	TTAGACTTCAGAGATTTATAGATTCCTCGTGAATAGGCTAAACTACTCTACAAATCTC	1086
QY	1053	GATGCTTTTGGTATTATGGACTATACCTTAACCTGGAAGAGTAGAGGATATACGATGAC	1112
Db	1087	GATGCTTTTGGTATTATGGACTATACCTTAACCTGGAAGAGTAGAGGATATACGATGAC	1146
QY	1113	ACCAACCGTATCAATACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAGATGCTAGCTAC	1172
Db	1147	ACCAACCGTATCAATACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAGATGCTAGCTAT	1206
QY	1173	CATTTAGCTGGTGGT 1187	
Db	1207	CAITTAGCCTATGAT 1221	

RESULT 10

US-07-854-596B-14
; Sequence 14, Application US/07854596B
; Patent No. 5434073
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith M
; APPLICANT: Hunter, Michael G
; APPLICANT: Czaplewski, Lloyd G
; TITLE OF INVENTION: Proteins and nucleic acids
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. John J. McDonnell
; STREET: Ten South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/854,596B
; FILING DATE: 03-JUN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,337
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEPEX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1335 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..1335
; OTHER INFORMATION: /note= "Streptokinase gene from S.
; OTHER INFORMATION: equisimilis"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 7..1326
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 7..1326
US-07-854-596B-14

Query Match 71.9%; Score 1108.6; DB 2; Length 1335;
Best Local Similarity 97.5%; Pred. No. 3.5e-309;
Matches 1126; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
Qy 33 ATAGCTGCTCCTGAATGGCTACTAGATCGTCTCTCTGTAATAACAGCCAAATTTGGTTGT 92
Db 85 ATTGCTGACCTTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCCAAATTTAGTTGT 144
Qy 93 AGCGTTGCTGTACTGTTGAGGGGACGAATCAAGACATAGTCTTAAATTTTTTGAATC 152
Db 145 AGCGTTGCTGTACTGTTGAGGGGACGAATCAAGACATAGTCTTAAATTTTTTGAAT 204
Qy 153 GATCTAAATCAACGACCTCTCATGAGGAAAGACAGACAGGCTTAAAGTCCAAATCA 212
Db 205 GACCTAAATCAACGACCTCTCATGAGGAAAGACAGACAGGCTTAAAGTCCAAATCA 264
Qy 213 AAACCAATTTGCTACTGATAGTGGCGGATGTCACTAAATTTGAGAAAGCTGCTTACTA 272
Db 265 AAACCAATTTGCTACTGATAGTGGCGGATGTCCACATAAATTTGAAAGCTGCTTACTA 324

Qy 273 AAGGCTATTCAAGAAACAAATTTGATCGCTAAAGTCCACAGTAACGACGACTACTTTGAGTC 332
Db 325 AAGGCTATTCAAGAAACAAATTTGATCGCTAAAGTCCACAGTAACGACGACTACTTTGAGTC 384
Qy 333 ATTGATTTTGAAGCGATGCAACATTTACTGATCGAAACGGCAAGGCTCTACTTTGCTGAC 392
Db 385 ATTGATTTTGAAGCGATGCAACATTTACTGATCGAAACGGCAAGGCTCTACTTTGCTGAC 444
Qy 393 AAAGATGGTTTGGTAACTTCCGACCCCAACCTGTCCAAAGAAATTTTGTAAAGCGGACAT 452
Db 445 AAAGATGGTTTGGTAACTTCCGACCCCAACCTGTCCAAAGAAATTTTGTAAAGCGGACAT 504
Qy 453 GTGCGGTTAGACCATATAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 512
Db 505 GTGCGGTTAGACCATATAAGAAACCAATACAAATCAAGCGAAATCTGTTGATGTG 564
Qy 513 GAATATCTGTACAGTTTACTTCCCTTAAACCCCTGATGACGATTTTCAAGACGAGTCTCAA 572
Db 565 GAATATCTGTACAGTTTACTTCCCTTAAACCCCTGATGACGATTTTCAAGACGAGTCTCAA 624
Qy 573 GATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTTACTA 632
Db 625 GATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTTACTA 684
Qy 633 GCTCAAGCACAAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTATGAACGTGAC 692
Db 685 GCTCAAGCACAAAGCATTTTAAACAAACCCATCCAGGCTATACGATTTATGAACGTGAC 744
Qy 693 TCCTCAATCGTCACTCATGACAAATGACATTTTCGGTACGATTTTACCAATGGATCAAGAG 752
Db 745 TCCTCAATCGTCACTCATGACAAATGACATTTTCGGTACGATTTTACCAATGGATCAAGAG 804
Qy 753 TTTACTTACCGTTTAAACCTCGGAAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 812
Db 805 TTTACTTACCATGTCAAAAATCGGAAACAAGCTTTATGAGATCAATAAAAAATCTGGTCTG 864
Qy 813 AATGAAGAAATAAACACACACGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 872
Db 865 AATGAAGAAATAAACACACACGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 924
Qy 873 GAAAGCCGTATGATCCCTTTGATCGCAGTCACCTTGAAACCTGTTCAACCATCAATACGTT 932
Db 925 GAAAGCCGTATGATCCCTTTGATCGCAGTCACCTTGAAACCTGTTCAACCATCAATACGTT 984
Qy 933 GATGTCGATACCAACGAATTTGCTAAAGTGAGCAGCTCTTAAACAGCTAGGAAACGTAAC 992
Db 985 GATGTCACACCAACGAATTTGCTAAAGTGAGCAGCTCTTAAACAGCTAGGAAACGTAAC 1044
Qy 993 TTAGACTTCAGAGATTTATACGATCCTCGGTGATAAGGCTAAACTCTTCAACAATCTC 1052
Db 1045 TTAGACTTCAGAGATTTATACGATCCTCGGTGATAAGGCTAAACTCTTCAACAATCTC 1104
Qy 1053 GATGCTTTTGGTATTATGAGTACTTAACTTGAAGAGTAGAGGATAATCAGATGAC 1112
Db 1105 GATGCTTTTGGTATTATGAGTACTTAACTTGAAGAGTAGAGGATAATCAGATGAC 1164
Qy 1113 ACCAAGCTATCAACCCGTTTATATGGGCAAGCAGCCGAGAGAGAGATGCTAGTAC 1172
Db 1165 ACCAAGCTATCAACCCGTTTATATGGGCAAGCAGCCGAGAGAGAGATGCTAGTAT 1224
Qy 1173 CATTTAGCTGGTGGT 1187
Db 1225 CATTTAGCTGGTGGT 1239

RESULT 11
US-07-854-596B-42
; Sequence 42, Application US/07854596B
; Patent No. 5434073
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith M
; APPLICANT: Hunter, Michael G
; APPLICANT: Czaplewski, Lloyd G

; TITLE OF INVENTION: Proteins and nucleic acids
 ; NUMBER OF SEQUENCES: 73
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: Dr. John J. McDonnell
 ; STREET: Ten South Wacker Drive, Suite 3000
 ; CITY: Chicago
 ; STATE: IL
 ; COUNTRY: USA
 ; ZIP: 60606
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/854,596B
 ; FILING DATE: 03-JUN-1992
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: McDonnell, John J
 ; REGISTRATION NUMBER: 26,949
 ; REFERENCE/DOCKET NUMBER: 92,337
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 312-715-1000
 ; TELEFAX: 312-715-1234
 ; TELEX: 910-221-5317
 ; INFORMATION FOR SEQ ID NO: 42:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1458 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cdna
 ; FEATURE:
 ; NAME/KEY: misc.feature
 ; LOCATION: 1..1458
 ; OTHER INFORMATION: /note= "Hirudin-streptokinase
 ; OTHER INFORMATION: fusion linked by Factor Xa cleavable IEGR"
 ;
 ; NAME/KEY: CDS
 ; LOCATION: 1..1449
 ;
 ; NAME/KEY: mat_peptide
 ; LOCATION: 1..1449
 ;
 ; US-07-854-596B-42

Query Match 71.9%; Score 1108.6; DB 2; Length 1458;
 Best Local Similarity 97.5%; Pred. No. 3.7e-309;
 Matches 1126; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
 QY 33 ATAGCTGGTCTCAATGGCTACTAGATCGTCTTCTGTAATACAGCCCAATGGTGGTT 92
 DB 208 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCCAATAGTGGTT 267
 QY 93 AGCGTTGCTGGTACTGTGTAGGGGACGAATCAAGACATTAGTCTTAATTTTGAATC 152
 DB 268 AGCGTTGCTGGTACTGTGTAGGGGACGAATCAAGACATTAGTCTTAATTTTGAATC 327
 QY 153 GATCTAACATCAGACCTGCTCATGGAGGAAAGACAGACGAGCTTAAAGTCCAAATCA 212
 DB 328 GACCTAACATCAGACCTGCTCATGGAGGAAAGACAGACGAGCTTAAAGTCCAAATCA 387
 QY 213 AAACCAATTTGCTACTGATAGTGGCGGATGTCATATAAATTGAGAAAGCTGACTTACTA 272
 DB 388 AAACCAATTTGCTACTGATAGTGGCGGATGTCACATATAAATTGAGAAAGCTGACTTACTA 447
 QY 273 AAGGCTATTCAAGAAATTTGATCGCTTAAGTCCAGATTAACGACGACTTTGAGTTC 332
 DB 448 AAGGCTATTCAAGAAATTTGATCGCTTAAGTCCAGATTAACGACGACTTTGAGTTC 507
 QY 333 ATTGATTTTGAAGCGATGCAACATTACTGTATCGAAACGCGAGGCTACTTTGCTGAC 392
 DB 508 ATTGATTTTGAAGCGATGCAACATTACTGTATCGAAACGCGAGGCTACTTTGCTGAC 567

QY 393 AAAAGATGGTTCGGTAAACCTTGGCCGACCCAACTGTCGAAGAAATTTTGTCTAAGCGGACAT 452
 DB 568 AAAAGATGGTTCGGTAAACCTTGGCCGACCCAACTGTCGAAGAAATTTTGTCTAAGCGGACAT 627
 QY 453 GTGCGCGTTAGACCATATAAAGAAAAACAATACAAAAACCAAGCGAAATCTGTTGATGTG 512
 DB 628 GTGCGCGTTAGACCATATAAAGAAAAACAATACAAAAACCAAGCGAAATCTGTTGATGTG 687
 QY 513 GAATATCTGTACAGTTTACTCCCTTAAACCTGTAGTACGATTTTACAGACCGGCTCTCAA 572
 DB 688 GAATATCTGTACAGTTTACTCCCTTAAACCTGTAGTACGATTTTACAGACCGGCTCTCAA 747
 QY 573 GATCTAAGCTATTGAAAACTAGCTATCGGTGACACCATCACATCTCAAGAAATTTACTA 632
 DB 748 GATCTAAGCTATTGAAAACTAGCTATCGGTGACACCATCACATCTCAAGAAATTTACTA 807
 QY 633 GCTCAAGCACAAGCAATTTTAAACAAAAACCAAGCGCTATACGATTTATGAACGTCAC 692
 DB 808 GCTCAAGCACAAGCAATTTTAAACAAAAACCAAGCGCTATACGATTTATGAACGTCAC 867
 QY 693 TCCTCAATCGTCACTATGACAAATGATATTTTCCGTACGATTTTACCAATGATCAAGAG 752
 DB 868 TCCTCAATCGTCACTATGACAAATGATATTTTCCGTACGATTTTACCAATGATCAAGAG 927
 QY 753 TTTTACTTACCGTGTAAATAATCGGGAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 812
 DB 928 TTTTACTTACCGTGTAAATAATCGGGAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 987
 QY 813 AATGAAGAAATAAACAACACATGACCTGATCTCTGAGAAATATTACGCTTTAAAAAAGGG 872
 DB 988 AATGAAGAAATAAACAACACATGACCTGATCTCTGAGAAATATTACGCTTTAAAAAAGGG 1047
 QY 873 GAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTTCAACCATCAATACGTT 932
 DB 1048 GAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTTCAACCATCAATACGTT 1107
 QY 933 GATGTCGATACCAAGCAATTTGCTAAAAGTGACGAGCTCTTAAACGCTAGGCAAGCTAAC 992
 DB 1108 GATGTCACACCAAGCAATTTGCTAAAAGTGACGAGCTCTTAAACGCTAGGCAAGCTAAC 1167
 QY 993 TTAGACTTCAGAGATTTATACGATCCTGCTGATAAGGCTAACTACTCTACAAATCTC 1052
 DB 1168 TTAGACTTCAGAGATTTATACGATCCTGCTGATAAGGCTAACTACTCTACAAATCTC 1227
 QY 1053 GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAGAGTAGAGGATAATCAGCATGAC 1112
 DB 1228 GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAGAGTAGAGGATAATCAGCATGAC 1287
 QY 1113 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAGGAGAGAAATGCTAGCTAC 1172
 DB 1288 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAGGAGAGAAATGCTAGCTAT 1347
 QY 1173 CATTTAGCTGGTGGT 1187
 DB 1348 CATTTAGCTTATGAT 1362

RESULT 12
 US-07-854-596B-46
 ; Sequence 46, Application US/07854596B
 ; Patent No. 5434073
 ; GENERAL INFORMATION:
 ; APPLICANT: Dawson, Keith M
 ; APPLICANT: Hunter, Michael G
 ; APPLICANT: Czaplewski, Lloyd G
 ; TITLE OF INVENTION: Proteins and nucleic acids
 ; NUMBER OF SEQUENCES: 73
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Dr. John J. McDonnell
 ; STREET: Ten South Wacker Drive, Suite 3000
 ; CITY: Chicago
 ; STATE: IL

COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/854,596B
FILING DATE: 03-JUN-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: McDonnell, John J
REGISTRATION NUMBER: 26,949
REFERENCE/DOCKET NUMBER: 92,337
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
TELEX: 910-221-5317
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 1467 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: misc.feature
LOCATION: 1..1467
OTHER INFORMATION: /note= "Streptokinase-hirudin
OTHER INFORMATION: fusion linked by Factor Xa-cleavable IEGR"
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1449
NAME/KEY: mat.peptide
LOCATION: 1..1449
US-07-854-596B-46

Query Match 71.9%; Score 1108.6; DB 2; Length 1467;

Best Local Similarity 97.5%; Pred. No. 3.7e-309;

Matches 1126; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY	33	ATAGCTGGTCCGTAATGGCTACTAGATCGTCTCTGTAATAACAGCCAAATGGTGT	92
DB	1	ATTGTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAAATAGTGT	60
QY	93	AGCGTTGCTGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC	152
DB	61	AGCGTTGCTGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAAT	120
QY	153	GATCTAACATACGACCTGCTCATGGAGGAAAGACAGACAGGCTTAAAGTCCAAATCA	212
DB	121	GACCTAACATACGACCTGCTCATGGAGGAAAGACAGACAGGCTTAAAGTCCAAATCA	180
QY	213	AAACCATTTGCTACTGATAGTGGCGGATGTCAATAAAGTGAAGAGCTGACTACTA	272
DB	181	AAACCATTTGCTACTGATAGTGGCGGATGCCACATAAAATTTGAAAGAGCTGACTA	240
QY	273	AAGGCTATTCAAGACAAATTTGATCGTTAAAGTCCAGTAAAGACGACTACTTTGAGTC	332
DB	241	AAGGCTATTCAAGACAAATTTGATCGTTAAAGTCCAGTAAAGACGACTACTTTGAGTC	300
QY	333	ATTGATTTTGAAGCGATCAACCAATTACTGATCGAAACGGCAAGGCTACTTTGCTGAC	392
DB	301	ATTGATTTTGAAGCGATCAACCAATTACTGATCGAAACGGCAAGGCTACTTTGCTGAC	360
QY	393	AAAGATGGTTCGGTAACCTTGGGACCCAACTGTCCAGAAATTTTGTCTAAGCGGACAT	452
DB	361	AAAGATGGTTCGGTAACCTTGGGACCCAACTGTCCAGAAATTTTGTCTAAGCGGACAT	420
QY	453	GTGCGGTTAGACCATATATAAGAAAAACCAATACAAAAACAGCGAAATCTGTTGATGTG	512

DB	421	GTGCGGTTAGACCATATATAAGAAAAACCAATACAAAAATCAAGCGAAATCTGTTGATGTG	480
QY	513	GAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAA	572
DB	481	GAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAA	540
QY	573	GATACTAAGCTATTGAAAAACACATAGCTTATCGGTGACACCATCATCTCAAGAATTACTA	632
DB	541	GATACTAAGCTATTGAAAAACACATAGCTTATCGGTGACACCATCATCTCAAGAATTACTA	600
QY	633	GCTCAAGCACAAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTATGAACGTCAC	692
DB	601	GCTCAAGCACAAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTATGAACGTCAC	660
QY	693	TCCTCAATCGTCACTCATGACAAATGACATTTTCGGTAGCATTTTACCAATGGATCAAGAG	752
DB	661	TCCTCAATCGTCACTCATGACAAATGACATTTTCGGTAGCATTTTACCAATGGATCAAGAG	720
QY	753	TTTACTTACCGTGTAAAAAATCGGAAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG	812
DB	721	TTTACTTACCATGTCAAAAAATCGGAAACAAGCTTTATGAGATCAATAAAAAATCTGGTCTG	780
QY	813	AATGAAGAAATAAAACAACACTGACCTGATCTCTGAGAAATATTACGCTCTTAAAAAAGGG	872
DB	781	AATGAAGAAATAAAACAACACTGACCTGATCTCTGAGAAATATTACGCTCTTAAAAAAGGG	840
QY	873	GAAGAGCGGTATGATCCCTTTGATCGGAGTCACCTTGAACCTGTTCAACATCAAAATAGCTT	932
DB	841	GAAGAGCGGTATGATCCCTTTGATCGGAGTCACCTTGAACCTGTTCAACATCAAAATAGCTT	900
QY	933	GATGTGCATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAAC	992
DB	901	GATGTCAACACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAAC	960
QY	993	TTAGACTTTCAGAGATTTTATACGATCTCTGCTGTAAGGCTAAAGCTACTCTACAAATCTC	1052
DB	961	TTAGACTTTCAGAGATTTTATACGATCTCTGCTGTAAGGCTAAAGCTACTCTACAAATCTC	1020
QY	1053	GATGCTTTTGGTATTATGACTATACCTTAACTGGAAAAAGTAGGAGATTAATCAGCATGAC	1112
DB	1021	GATGCTTTTGGTATTATGACTATACCTTAACTGGAAAAAGTAGGAGATTAATCAGCATGAC	1080
QY	1113	ACCAACCGTATACACCGTTTATATGGGCAAGCGCCCGAAGGAGAGATGCTAGCTAC	1172
DB	1081	ACCAACCGTATACACCGTTTATATGGGCAAGCGCCCGAAGGAGAGATGCTAGCTAC	1140
QY	1173	CATTTAGCTGGTGT 1187	
DB	1141	CATTTAGCTATGAT 1155	

RESULT 13

US-07-854-596B-34

; Sequence 34, Application US/07854596B

; Patent No. 5434073

; GENERAL INFORMATION:

; APPLICANT: Dawson, Keith M

; APPLICANT: Hunter, Michael G

; APPLICANT: Czaplewski, Lloyd G

; TITLE OF INVENTION: Proteins and nucleic acids

; NUMBER OF SEQUENCES: 73

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Dr. John J. McDonnell

; STREET: Ten South Wacker Drive, Suite 3000

; CITY: Chicago

; STATE: IL

; COUNTRY: USA

; ZIP: 60606

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/854,596B
 ; FILING DATE: 03-JUN-1992
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: McDonnell, John J
 ; REGISTRATION NUMBER: 26,949
 ; REFERENCE/DOCKET NUMBER: 92,337
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 312-715-1000
 ; TELEFAX: 312-715-1234
 ; TELEX: 910-221-5317
 ; INFORMATION FOR SEQ ID NO: 34:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2589 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: 1..2589
 ; OTHER INFORMATION: /note=
 ; OTHER INFORMATION: "OmpA-Streptokinase-fusion linked
 ; OTHER INFORMATION: by thrombin-cleavable VELOGVPRG"
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 4..2580
 ; FEATURE:
 ; NAME/KEY: mat_peptide
 ; LOCATION: 4..2580
 ;
 ; US-07-854-596B-34

Query Match 71.9%; Score 1108.6; DB 2; Length 2589;
 Best Local Similarity 97.5%; Pred. No. 4.7e-309;
 Matches 1126; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

 QY 33 ATAGCTGGTCCCTGAATGGCTACTAGATCGTCTTCTGTAATAACAGCCAAATGGTGT 92
 DB 67 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAGCCAAATAGTGT 126
 QY 93 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 152
 DB 127 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAAT 186
 QY 153 GATCTAATCATCAGCCTGCTATGAGGAGGAGCAAGCAAGCTTAAGTCCAAATCA 212
 DB 187 GACCTAATCATCAGCCTGCTATGAGGAGGAGCAAGCAAGCTTAAGTCCAAATCA 246
 QY 213 AAACCATTTGCTACTGATAGTGGCGGATCTCATATAAACTTCAGAAAGCTGACTTACTA 272
 DB 247 AAACCATTTGCTACTGATAGTGGCGGATCTCATATAAACTTCAGAAAGCTGACTTACTA 306
 QY 273 AAGGCTATTCAAGAACAAATTGATCGCTTAACGCTCCACAGTAACGACGACTTCTTGAGTC 332
 DB 307 AAGGCTATTCAAGAACAAATTGATCGCTTAACGCTCCACAGTAACGACGACTTCTTGAGTC 366
 QY 333 ATTGATTTTGAAGCGGATCAACCAATTAAGTTCGAAACGGCAAGGCTCTACTTTGCTGAC 392
 DB 367 ATTGATTTTGAAGCGGATCAACCAATTAAGTTCGAAACGGCAAGGCTCTACTTTGCTGAC 426
 QY 393 AAGATAGTTCGGTACCTTGGGACCAACCTGTCAGAAATTTTTCGTAAGCGGACAT 452
 DB 427 AAGATAGTTCGGTACCTTGGGACCAACCTGTCAGAAATTTTTCGTAAGCGGACAT 486
 QY 453 GTGCGGTTAGACCATATATAAGAAAAACCAATACAAAAACCAAGCAATCTGTTGATGTG 512
 DB 487 GTGCGGTTAGACCATATATAAGAAAAACCAATACAAAAACCAAGCAATCTGTTGATGTG 546
 QY 513 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAA 572
 DB 547 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAA 606

QY 573 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCAATCTCAAGAATTACTA 632
 DB 607 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCAATCTCAAGAATTACTA 666
 QY 633 GCTCAAGCACAAGACATTTTAAACAAAAACCAACCCAGGCTATACGATTTATGAACGTGAC 692
 DB 667 GCTCAAGCACAAGACATTTTAAACAAAAACCAACCCAGGCTATACGATTTATGAACGTGAC 726
 QY 693 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTAGATTTTACCAATGGATCAAGAG 752
 DB 727 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTAGATTTTACCAATGGATCAAGAG 786
 QY 753 TTTACTTTACCGTGTAAAAAATCGGGAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 812
 DB 787 TTTACTTTACCATGTCAAAAATCGGGAACAAGCTTTATGAGATCAATAAAAAATCTGGTCTG 846
 QY 813 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAAGTCTCTTAAAAAAGGG 872
 DB 847 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAAGTCTCTTAAAAAAGGG 906
 QY 873 GAAAAGCCGTATGATCCCTTTGATCGGAGTCACTTGAACCTGTTCAACCATCAATACGTT 932
 DB 907 GAAAAGCCGTATGATCCCTTTGATCGGAGTCACTTGAACCTGTTCAACCATCAATACGTT 966
 QY 933 GATGTCGATACCAACGAATGCTAAAAAGTGAGCAGCTCTTTAAACAGCTAGCGAAACGTAAC 992
 DB 967 GATGTCGATACCAACGAATGCTAAAAAGTGAGCAGCTCTTTAAACAGCTAGCGAAACGTAAC 1026
 QY 993 TTAGACTTCAGAGATTTATAGATCTCTGCTGATTAAGGCTAAACTTCTTACAAACATCTC 1052
 DB 1027 TTAGACTTCAGAGATTTATAGATCTCTGCTGATTAAGGCTAAACTTCTTACAAACATCTC 1086
 QY 1053 GATGCTTTTGGTATTATGAGTACTTAACTTAACTGGAAGCTAGAGGATTAATCAGATGAC 1112
 DB 1087 GATGCTTTTGGTATTATGAGTACTTAACTTAACTGGAAGCTAGAGGATTAATCAGATGAC 1146
 QY 1113 ACCAACCGTATCAACCGTTTATATGGGAAAGGACCCGAAAGGAGAGATGCTAGCTAC 1172
 DB 1147 ACCAACCGTATCAACCGTTTATATGGGAAAGGACCCGAAAGGAGAGATGCTAGCTAC 1206
 QY 1173 CATTTAGCTGGTGGT 1187
 DB 1207 CATTTAGCTGGTGGT 1221

RESULT 14
 US-09-374-038-11
 ; Sequence 11, Application US/09374038
 ; Patent No. 6309873
 ; GENERAL INFORMATION:
 ; APPLICANT: Madrazo, Isis Del Carmen Torrens
 ; APPLICANT: Garcia, Jose De Jesus De La Fuente
 ; APPLICANT: Ojalvo, Ariana Garcia
 ; APPLICANT: Menendez, Alina Seralea
 ; APPLICANT: Escalona, Elder Pupo
 ; APPLICANT: Masso, Julio Raul Fernandez
 ; APPLICANT: Griego, Martha De Jesus Gonzalez
 ; TITLE OF INVENTION: STREPTOKINASE MUTANTS
 ; FILE REFERENCE: Sequence Listings 1-14 re: 976-5
 ; Patent No. 6309873
 ; CURRENT APPLICATION NUMBER: US/09/374,038
 ; CURRENT FILING DATE: 1999-08-13
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 11
 ; LENGTH: 1209
 ; TYPE: DNA
 ; ORGANISM: Streptococcus equisimilis
 ; US-09-374-038-11

Query Match 70.4%; Score 1084.4; DB 3; Length 1209;
 Best Local Similarity 98.1%; Pred. No. 3.2e-302;
 Matches 1097; Conservative 0; Mismatches 21; Indels 0; Gaps 0;


```
QY 70 TAAATAACAGCCAAATGGTTGTTAGCGTTGCTGCTACTGATCATGAGAGGAAAGACAG 129
Db 2 TGAACAACAGCCAAATGGTTGTTAGCGTTGCTGCTACTGATCATGAGAGGAAAGACAG 61

QY 130 TTAGTCTTAAATTTTGGAAATCGATCTAACATCAGACCTGCTCATGAGAGGAAAGACAG 189
Db 62 TTAGTCTTAAATTTTGGAAATCGATCTAACATCAGACCTGCTCATGAGAGGAAAGACAG 121

QY 190 ACGAAGGCTTAAGTCCAAATCAAAACCAATTTGCTACTGATGATGCGGCGATGTCACATA 249
Db 122 ACGAAGGCTTAAGTCCAAATCAAAACCAATTTGCTACTGATGATGCGGCGATGTCACATA 181

QY 250 AACTTTGAGAAAGCTGACTTACTAAAGGCTATTCAAGAAACAATTTGATCGCTAAAGCTCACA 309
Db 182 AACTTTGAGAAAGCTGACTTACTAAAGGCTATTCAAGAAACAATTTGATCGCTAAAGCTCACA 241

QY 310 GTAACGACGACTACTTTGAGGCTCATTTGTTGCAAGCGATGCAACCATTTACTGATCGAA 369
Db 242 GTAACGACGACTACTTTGAGGCTCATTTGTTGCAAGCGATGCAACCATTTACTGATCGAA 301

QY 370 ACGGCAAGGCTACTTTGCTGACAAAGATGTTTCGGTAACTTTGCCGACCAACCTGTCC 429
Db 302 ACGGCAAGGCTACTTTGCTGACAAAGATGTTTCGGTAACTTTGCCGACCAACCTGTCC 361

QY 430 AAGAATTTTGTCTAAGCGGACATGTCGGGTTAGACCATATAAAGAAACCAATACAAA 489
Db 362 AAGAATTTTGTCTAAGCGGACATGTCGGGTTAGACCATATAAAGAAACCAATACAAA 421

QY 490 ACCAAGCGAAATCTGTTGATGGAATATACGTACAGATTTACTCCCTTTAAACCCCTGATG 549
Db 422 ATCAAGCGAAATCTGTTGATGGAATATACGTACAGATTTACTCCCTTTAAACCCCTGATG 481

QY 550 AGATTTTCAGACAGGCTCTCAAGATACCTAAGCTATTGAAACACTAGCTATCGGTGACA 609
Db 482 AGATTTTCAGACAGGCTCTCAAGATACCTAAGCTATTGAAACACTAGCTATCGGTGACA 541

QY 610 CCATCACATCTCAAGAAATTTACTAGCTCAAGGCAAAAGCAATTTTAAACCAAAACCCAG 669
Db 542 CCATCACATCTCAAGAAATTTACTAGCTCAAGGCAAAAGCAATTTTAAACCAAAACCCAG 601

QY 670 GCTATACGATTTATGAACGTGACTCCTCAATCGTCACTCATGACAAATGACATTTTCGGTA 729
Db 602 GCTATACGATTTATGAACGTGACTCCTCAATCGTCACTCATGACAAATGACATTTTCGGTA 661

QY 730 CGATTTTACCAATGATCAAGAGTTTACTTACCGTGTAAATTCGGGAAACAAGCTTATA 789
Db 662 CGATTTTACCAATGATCAAGAGTTTACTTACCGTGTAAATTCGGGAAACAAGCTTATA 721

QY 790 GGATCAATAAATAATCTGCTGTAATGAAGAAATAAACCACTGACCTGATCTCTGGA 849
Db 722 AGATCAATAAATAATCTGCTGTAATGAAGAAATAAACCACTGACCTGATCTCTGGA 781

QY 850 AATATTAAGTCTTAAAGGAGGAAAGCCGATGATCCCTTTGATCGCAGTCACTTGA 909
Db 782 AATATTAAGTCTTAAAGGAGGAAAGCCGATGATCCCTTTGATCGCAGTCACTTGA 841

QY 910 AACTGTTTCAACATCAATAACGTTGATGTCGATACCAACGAATGCTTAAAGGTGAGCAGC 969
Db 842 AACTGTTTCAACATCAATAACGTTGATGTCGATACCAACGAATGCTTAAAGGTGAGCAGC 901

QY 970 TCCTTAACAGCTAGCGAAGCTTAAGCTTACAGATTTTATACGATCTCTCGGTAAAG 1029
Db 902 TCCTTAACAGCTAGCGAAGCTTAAGCTTACAGATTTTATACGATCTCTCGGTAAAG 961

QY 1030 CTAAACTACTCTACAACTCTCGATGCTTTTGGTATTATGGAATATACCTTAACCTGGA 1089
Db 962 CTAAACTACTCTACAACTCTCGATGCTTTTGGTATTATGGAATATACCTTAACCTGGA 1021

QY 1090 AAGTAGAGGATAATCAGATGACCAACCGTATCATACCGTTTATATGGGCAAGCGAC 1149
Db 1022 AAGTAGAGGATAATCAGATGACCAACCGTATCATACCGTTTATATGGGCAAGCGAC 1081
```

```
QY 1150 CCGAAGGAGAGAATGCTAGCTACCATTTAGCTGGTGGT 1187
Db 1082 CCGAAGGAGAGAATGCTAGCTATCATTTAGSCTTGAT 1119

RESULT 15
US-09-658-179-11
; Sequence 11, Application US/09658179
; Patent No. 6413759
; GENERAL INFORMATION:
; APPLICANT: Madrazo, Isis Del Carmen Torrens
; APPLICANT: Garcia, Jose De Jesus De La Fuente
; APPLICANT: Ojalvo, Ariana Garcia
; APPLICANT: Menendez, Alina Seralena
; APPLICANT: Escalona, Elder Pupo
; APPLICANT: Masso, Julio Raul Fernandez
; APPLICANT: Griego, Marcha De Jesus Gonzalez
; TITLE OF INVENTION: STREPTOKINASE MUTANTS
; FILE REFERENCE: Sequence Listings 1-14 re: 976-5
; Patent No. 6413759
; CURRENT APPLICATION NUMBER: US/09/658,179
; CURRENT FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 1209
; TYPE: DNA
; ORGANISM: Streptococcus equisimilis
US-09-658-179-11

Query Match 70.4%; Score 1084.4; DB 3; Length 1209;
Best Local Similarity 98.1%; Pred. No. 3.2e-302;
Matches 1097; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 70 TAAATAACAGCCAAATGGTTGTTAGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACA 129
Db 2 TGAACAACAGCCAAATGGTTGTTAGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACA 61

QY 130 TTAGTCTTAAATTTTGGAAATCGATCTAACATCAGACCTGCTCATGAGAGGAAAGACAG 189
Db 62 TTAGTCTTAAATTTTGGAAATCGATCTAACATCAGACCTGCTCATGAGAGGAAAGACAG 121

QY 190 AGCAAGGCTTAAAGTCCAAATCAAAACCAATTTGCTACTGATGATGCGGCGATGTCACATA 249
Db 122 AGCAAGGCTTAAAGTCCAAATCAAAACCAATTTGCTACTGATGATGCGGCGATGTCACATA 181

QY 250 AACTTTGAGAAAGCTGACTTACTAAAGGCTATTCAAGAAACAATTTGATCGCTAAAGCTCACA 309
Db 182 AACTTTGAGAAAGCTGACTTACTAAAGGCTATTCAAGAAACAATTTGATCGCTAAAGCTCACA 241

QY 310 GTAACGACGACTACTTTGAGGCTCATTTGTTGCAAGCGATGCAACCATTTACTGATCGAA 369
Db 242 GTAACGACGACTACTTTGAGGCTCATTTGTTGCAAGCGATGCAACCATTTACTGATCGAA 301

QY 370 ACGGCAAGGCTACTTTGCTGACAAAGATGTTTCGGTAACTTTGCCGACCAACCTGTCC 429
Db 302 ACGGCAAGGCTACTTTGCTGACAAAGATGTTTCGGTAACTTTGCCGACCAACCTGTCC 361

QY 430 AAGAATTTTGTCTAAGCGGACATGTCGGGTTAGACCATATAAAGAAACCAATACAAA 489
Db 362 AAGAATTTTGTCTAAGCGGACATGTCGGGTTAGACCATATAAAGAAACCAATACAAA 421

QY 490 ACCAAGCGAAATCTGTTGATGGAATATACGTACAGATTTACTCCCTTTAAACCCCTGATG 549
Db 422 ATCAAGCGAAATCTGTTGATGGAATATACGTACAGATTTACTCCCTTTAAACCCCTGATG 481

QY 550 AGATTTTCAGACAGGCTCTCAAGATACCTAAGCTATTGAAACACTAGCTATCGGTGACA 609
Db 482 AGATTTTCAGACAGGCTCTCAAGATACCTAAGCTATTGAAACACTAGCTATCGGTGACA 541

QY 610 CCATCACATCTCAAGAAATTTACTAGCTCAAGGCAAAAGCAATTTTAAACCAAAACCCAG 669
Db 542 CCATCACATCTCAAGAAATTTACTAGCTCAAGGCAAAAGCAATTTTAAACCAAAACCCAG 601
```

Qy	670	GCTATACGATTATGAA	CGTGACTCCTCAATCGTCACTCATGACAATGACATTTTCGGTA	729
Db	602	GCTATACGATTATGAA	CGTGACTCCTCAATCGTCACTCATGACAATGACATTTTCGGTA	661
Qy	730	CGATTTTACCAATGGATCA	AGAGTTTACTTACCGTGTAAATAATCGGGAAACAAGCTTATA	789
Db	662	CGATTTTACCAATGGATCA	AGAGTTTACTTACCATGTCAAAAATCGGGAAACAAGCTTATG	721
Qy	790	GGATCAATAAAAATCTGGT	CTGAATGAAGAAATAAACACACTGACTGATCTCTGAGA	849
Db	722	AGATCAATAAAAATCTGGT	CTGAATGAAGAAATAAACACACTGACTGATCTCTGAGA	781
Qy	850	AATATTACGTCCTTAAAA	AGGGGAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGA	909
Db	782	AATATTACGTCCTTAAAA	AGGGGAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGA	841
Qy	910	AACGTTCACCATCAAAAT	ACGTTGATGTCGATACCAACGAATTCGCTAAAAAGTCAGCAGC	969
Db	842	AACGTTCACCATCAAAAT	ACGTTGATGTCGATACCAACGAATTCGCTAAAAAGTCAGCAGC	901
Qy	970	TCTTAACAGCTAGCGAC	TAACTTAGACTTCAGAGATTTATACGATCCTCGTGATAAGG	1029
Db	902	TCTTAACAGCTAGCGAC	TAACTTAGACTTCAGAGATTTATACGATCCTCGTGATAAGG	961
Qy	1030	CTAAACTACTCTACAACA	ATCTCGATGCTTTTGGTATTATGGACTATACCTTAACTGAAA	1089
Db	962	CTAAACTACTCTACAACA	ATCTCGATGCTTTTGGTATTATGGACTATACCTTAACTGAAA	1021
Qy	1090	AAGTAGAGGATAATCAG	ATGACACCAACCGTATCATACCGTTTATATGGGCAAGCGAC	1149
Db	1022	AAGTAGAGGATAATCAG	ATGACACCAACCGTATCATACCGTTTATATGGGCAAGCGAC	1081
Qy	1150	CCGAGGAGAGAATGCTA	GTACCATTTAGCTGGTGGT	1187
Db	1082	CCGAGGAGAGAATGCTA	GTATCATTTAGCCTATGAT	1119

Search completed: February 1, 2006, 12:43:41
 Job time : 273.069 secs

1
Pending Nucleic Acid and Pending Amino Acid database searches generate two sets of results each. The Pending databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches. Searches run against the Nucleic Acid Pending database produce two sets of results, with the extensions **.rnpm** and **.rnpn**. Searches run against the Amino Acid Pending database produce two sets of results, with the extensions **.rapm** and **.rapn**.

Because they contain data that is confidential, the results of Pending database searches should not be left in the case .

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2006, 17:06:30 ; Search time 1273.28 Seconds
(without alignments)
10008.068 Million cell updates/sec

Title: US-09-940-235-9

Perfect score: 1541

Sequence: 1 ttgtttaaactttaagaagg.....ccttcacgatgttcttag 1541

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA Main:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1541	100.0	1541	3	US-09-940-235-9 Sequence 9, Appli
2	1496.2	97.1	2096	3	US-09-940-235-12 Sequence 12, Appl
3	1185	76.9	1661	3	US-09-940-235-10 Sequence 10, Appl
4	1179	76.5	1327	3	US-09-940-235-6 Sequence 6, Appli
5	1166.2	75.7	1377	3	US-09-940-235-5 Sequence 5, Appli
6	1134.2	73.6	1782	3	US-09-940-235-11 Sequence 11, Appl
7	1132.6	73.5	1245	3	US-09-940-235-1 Sequence 1, Appli
8	993.4	64.5	1323	8	US-10-474-792-657 Sequence 657, App
9	330	21.4	777	3	US-09-940-235-3 Sequence 3, Appli
10	330	21.4	2127	6	US-10-210-120-49 Sequence 49, Appl
11	330	21.4	2127	9	US-10-909-035-49 Sequence 49, Appl
12	330	21.4	2127	9	US-10-909-035-49 Sequence 49, Appl
13	330	21.4	2443	8	US-10-741-601-70 Sequence 70, Appl
14	330	21.4	2443	8	US-10-741-600-238 Sequence 238, App
15	330	21.4	2488	8	US-10-741-601-75 Sequence 75, Appl
16	330	21.4	2488	8	US-10-741-600-244 Sequence 244, App
17	330	21.4	4295	6	US-10-144-194A-51 Sequence 51, Appl
18	330	21.4	4295	8	US-10-491-566-51 Sequence 51, Appl
19	330	21.4	6510	8	US-10-741-601-72 Sequence 72, Appl
20	330	21.4	6510	8	US-10-741-600-241 Sequence 241, App
21	330	21.4	6988	7	US-10-236-392-1 Sequence 1, Appli
22	330	21.4	7361	7	US-10-236-392-3 Sequence 3, Appli
23	330	21.4	7677	9	US-10-956-157-4995 Sequence 4995, Ap

ALIGNMENTS

RESULT 1
US-09-940-235-9
; Sequence 9, Application US/09940235
; Publication No. US20030059921A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammar
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/09/940,235
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1541
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hybrid cassette
US-09-940-235-9

Query Match 100.0%; Score 1541; DB 3; Length 1541;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1541; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTGTGTTAACTTTAAGAAGGAGATATACCATGATAGCTGCTCTGTAATGCTACTAGATC 60
DB 1 TTGTGTTAACTTTAAGAAGGAGATATACCATGATAGCTGCTCTGTAATGCTACTAGATC 60
QY 61 GTCTTTCTGTAATAACAGCAATGCTGTTAGCGTGTGCTGTTAGCGGACGA 120
DB 61 GTCTTTCTGTAATAACAGCAATGCTGTTAGCGTGTGCTGTTAGCGGACGA 120
QY 121 ATCAAGACATAGCTTTAAATTTTTTGAATCGATCTAACATCAGCACCTGCTCATGGAG 180

Sequence 38, Appl
Sequence 574, Appl
Sequence 63, Appl
Sequence 69, Appl
Sequence 75, Appl
Sequence 1289, Ap
Sequence 222, App
Sequence 79, Appl
Sequence 654, App
Sequence 654, App
Sequence 88, Appl
Sequence 603, App
Sequence 5877, Ap
Sequence 52, Appl
Sequence 2, Appl
Sequence 77, Appl
Sequence 245, App
Sequence 78, Appl
Sequence 246, App
Sequence 6, Appli

```
Db 121 ATCAAGACATTAGTCTTAAATTTTGGAAATCGATCTAAACATCAGCACCTGCTCATGGAG 180
Qy 181 GAAAGACAGACGAGCGCTTAAGTCCAAATCAAAACCAATTTGCTACTGATAGTGGCGGA 240
Db 181 GAAAGACAGACGAGCGCTTAAGTCCAAATCAAAACCAATTTGCTACTGATAGTGGCGGA 240
Qy 241 TGTCACATAAACTTGAGAAAGCTGACTTAAAGGCTATTCAAGAAACAATTCATCGCTA 300
Db 241 TGTCACATAAACTTGAGAAAGCTGACTTAAAGGCTATTCAAGAAACAATTCATCGCTA 300
Qy 301 AGCTCCACAGTAACGACGACTACTTTGAGGTCAATTTGCAAGCGATGCAACCATTA 360
Db 301 AGCTCCACAGTAACGACGACTACTTTGAGGTCAATTTGCAAGCGATGCAACCATTA 360
Qy 361 CTGATCGAAACCGCAAGGTCTACTTTGCTGCAAAAGATGGTTGCGTAACTTCCGACCC 420
Db 361 CTGATCGAAACCGCAAGGTCTACTTTGCTGCAAAAGATGGTTGCGTAACTTCCGACCC 420
Qy 421 AACCTGTCCAGAAATTTTGTGTAAGCGGACATGTGCGGTTAGACCATATAAGAAAAAAC 480
Db 421 AACCTGTCCAGAAATTTTGTGTAAGCGGACATGTGCGGTTAGACCATATAAGAAAAAAC 480
Qy 481 CAATACAAAACCAAGCGAAATCTGTTGATGTGGAATATATCTGTACAGTTTACTCCCTTAA 540
Db 481 CAATACAAAACCAAGCGAAATCTGTTGATGTGGAATATATCTGTACAGTTTACTCCCTTAA 540
Qy 541 ACCCTGATGACGATTTACAGACAGGTCTCAAGATACTAAGCTATTGAAACACTAGCTA 600
Db 541 ACCCTGATGACGATTTACAGACAGGTCTCAAGATACTAAGCTATTGAAACACTAGCTA 600
Qy 601 TCGGTGACACCATCATCTCAAGAAATTAAGTCAAGCACAAGCAATTTTAAACAAA 660
Db 601 TCGGTGACACCATCATCTCAAGAAATTAAGTCAAGCACAAGCAATTTTAAACAAA 660
Qy 661 ACCCTCCAGGCTATACGATTTATGAACGTGACTCTCAATCGTCACTCATGACAAATGACA 720
Db 661 ACCCTCCAGGCTATACGATTTATGAACGTGACTCTCAATCGTCACTCATGACAAATGACA 720
Qy 721 TTTTCCGTACGATTTTACCAATGGATCAAGAGTTTACTTACCGTGTAAATAACCGGAAAC 780
Db 721 TTTTCCGTACGATTTTACCAATGGATCAAGAGTTTACTTACCGTGTAAATAACCGGAAAC 780
Qy 781 AAGCTTATAGGATCAATAAAATCTGGTCTGAATGAAGAAATAAACAACACTGACTGA 840
Db 781 AAGCTTATAGGATCAATAAAATCTGGTCTGAATGAAGAAATAAACAACACTGACTGA 840
Qy 841 TCTCTGAGAAATATTACGTCCTTAAATAAGGGGAAAGCGTATGATCCCTTTGATCGCA 900
Db 841 TCTCTGAGAAATATTACGTCCTTAAATAAGGGGAAAGCGTATGATCCCTTTGATCGCA 900
Qy 901 GTCACTTGAACCTGTTACCATCAATACTGTTGATGTCGATACCAAGAAATGCTTAAAAA 960
Db 901 GTCACTTGAACCTGTTACCATCAATACTGTTGATGTCGATACCAAGAAATGCTTAAAAA 960
Qy 961 GTGACGAGCTTTAAGCTAGCGAAGCTTAACCTTAGACTTCAGAGATTTATAGCATCCTC 1020
Db 961 GTGACGAGCTTTAAGCTAGCGAAGCTTAACCTTAGACTTCAGAGATTTATAGCATCCTC 1020
Qy 1021 GTGATAAGGCTAAATCTACTCTACAACTCTCGATGCTTTTGGTATTATGGACTATACCT 1080
Db 1021 GTGATAAGGCTAAATCTACTCTACAACTCTCGATGCTTTTGGTATTATGGACTATACCT 1080
Qy 1081 TAACTGGAAGTAGAGGATAATACGATGACACCAACCGTATCATAACCGTTTATATGG 1140
Db 1081 TAACTGGAAGTAGAGGATAATACGATGACACCAACCGTATCATAACCGTTTATATGG 1140
Qy 1141 GCAAGCGACCGAAGAGAGATGCTAGCTACATTTAGCTGTGTGGCGGCGCAAC 1200
Db 1141 GCAAGCGACCGAAGAGAGATGCTAGCTACATTTAGCTGTGTGGCGGCGCAAC 1200
Qy 1201 AGATTGTACCATAGCTGAGAGTGTGTTGATCATGCTGCTGGGACTTCTCTATGCTGTCG 1260
```

```
Db 1201 AGATTGTACCATAGCTGAGAGTGTGTTGATCATGCTGCTGGGACTTCTATGTTGTCG 1260
Qy 1261 GAGAAACGTGGGAGAGCCCTACCAAGGCTGGGATGATGATGTTGCTTTCCTCGGAG 1320
Db 1261 GAGAAACGTGGGAGAGCCCTACCAAGGCTGGGATGATGATGTTGCTTTCCTCGGAG 1320
Qy 1321 AAGCGAGCGGACGCATCTGACCTTCTAGAAATAGATCAACGATCAGGACACAAGGA 1380
Db 1321 AAGCGAGCGGACGCATCTGACCTTCTAGAAATAGATCAACGATCAGGACACAAGGA 1380
Qy 1381 CATCTATAGAAATGAGACACCTGGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGT 1440
Db 1381 CATCTATAGAAATGAGACACCTGGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGT 1440
Qy 1441 GCATCTCCACAGCAACGCGGAGGAGTGGAGTGTGAGAGCACACCTCTGTGAGA 1500
Db 1441 GCATCTCCACAGCAACGCGGAGGAGTGGAGTGTGAGAGCACACCTCTGTGAGA 1500
Qy 1501 CCACATCGAGCGATCTGGCCCTTACCCGATGTTGCTTAG 1541
Db 1501 CCACATCGAGCGATCTGGCCCTTACCCGATGTTGCTTAG 1541
```

RESULT 2

```
US-09-940-235-12
; Sequence 12, Application US/09940235
; Publication No. US20030059921A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammar
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/09/940,235
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2096
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hybrid cassette
US-09-940-235-12
```

```
Query Match 97.1%; Score 1496.2; DB 3; Length 2096;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1501; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
```

```
Qy 33 ATAGCTGGTCTCAATGGCTACTAGATCGTCTCTGTAATAATACAGCCCAATTTGGTT 92
Db 588 ATTGCTGGACCTGAGTGGCTGTAGACCGCTCCATCTGTCAACCAACAGCCCAATTTGGTT 647
Qy 93 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 152
Db 648 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 707
Qy 153 GATCTAAATCAACGACCTGCTCATGGAGGAAAGACAGCAAGGCTTAAGTCCAAATCA 212
Db 708 GATCTAAATCAACGACCTGCTCATGGAGGAAAGACAGCAAGGCTTAAGTCCAAATCA 767
Qy 213 AAACCATTTGCTACTGATGAGTGGCGGATGTCATATAAATTTGAGAAAGCTGACTTACTA 272
```



```
Db 391 TGTTCATATAAATTGAGAAAGCTGCTTACTTAAAGGCTATTCAAGAACCAATTGATCGCTA 450
Qy 301 AGCTCCACAGTAACGACGACTACTTTGAGGTCAATGATTTTGCAGGCGATGCAACCAATTA 360
Db 451 AGCTCCACAGTAACGACGACTACTTTGAGGTCAATGATTTTGCAGGCGATGCAACCAATTA 510
Qy 361 CTGATCGAAACCGCAAGGCTCTACTTTGCTGACAAAGATGTTCCGTAACCTTGCCGACCC 420
Db 511 CTGATCGAAACCGCAAGGCTCTACTTTGCTGACAAAGATGTTCCGTAACCTTGCCGACCC 570
Qy 421 AACCTGTCCAAGAAATTTTGTGTAAGCGGACATGTGCGCGTTAGACCATATAAAGAAAC 480
Db 571 AACCTGTCCAAGAAATTTTGTGTAAGCGGACATGTGCGCGTTAGACCATATAAAGAAAC 630
Qy 481 CAATACAAACCAAGCGAAATCTGTTGATGTGGAATATCTGTACAGTTTACTCCCTTAA 540
Db 631 CAATACAAACCAAGCGAAATCTGTTGATGTGGAATATCTGTACAGTTTACTCCCTTAA 690
Qy 541 ACCCTGATGACGATTTACAGCAGGCTCTCAAGATACTAAGCTATTGAAACCACTAGCTA 600
Db 691 ACCCTGATGACGATTTACAGCAGGCTCTCAAGATACTAAGCTATTGAAACCACTAGCTA 750
Qy 601 TCGGTGACACCATCATCTCAAGAAATTAAGTCTCAAGCACAAGCAATTTTAAACAAA 660
Db 751 TCGGTGACACCATCATCTCAAGAAATTAAGTCTCAAGCACAAGCAATTTTAAACAAA 810
Qy 661 ACCACCCAGGCTATAGGATTTATGAACGTCATCTCAATGTCATCTCATGACATGACATA 720
Db 811 ACCACCCAGGCTATAGGATTTATGAACGTCATCTCAATGTCATCTCATGACATGACATA 870
Qy 721 TTTTCCGTCAGATTTTACCAATGATCAAGAGTTTACTTACCGTGTAAAGATTCGGGAAC 780
Db 871 TTTTCCGTCAGATTTTACCAATGATCAAGAGTTTACTTACCGTGTAAAGATTCGGGAAC 930
Qy 781 AAGCTTTATAGGATCAATAAATAATCTGGTCTGAATGAAGAAATFAAACACACTGACCTGA 840
Db 931 AAGCTTTATAGGATCAATAAATAATCTGGTCTGAATGAAGAAATFAAACACACTGACCTGA 990
Qy 841 TCTCTGAGAAATATTAGTCCTTAAAGAAAGGGGAAAGCGGTATGATTCCTTTGATCGCA 900
Db 991 TCTCTGAGAAATATTAGTCCTTAAAGAAAGGGGAAAGCGGTATGATTCCTTTGATCGCA 1050
Qy 901 GTCACCTGAACTGTTCCACATCAATAGCTTTGATGTCGATACCAACGAATTGCTTAAAA 960
Db 1051 GTCACCTGAACTGTTCCACATCAATAGCTTTGATGTCGATACCAACGAATTGCTTAAAA 1110
Qy 961 GTGACGAGCTCTTAACAGCTAGCGAACGTAACCTTAGACTTTCAGAGATTTTATACGATCCTC 1020
Db 1111 GTGACGAGCTCTTAACAGCTAGCGAACGTAACCTTAGACTTTCAGAGATTTTATACGATCCTC 1170
Qy 1021 GTGATAAGGCTAAAATCTCTCAACAAATCTCGATGCTTTTGGTATTATGGAATATACCT 1080
Db 1171 GTGATAAGGCTAAAATCTCTCAACAAATCTCGATGCTTTTGGTATTATGGAATATACCT 1230
Qy 1081 TAACTGGAAAGTAGAGGATTAATCAGATGACACCAACCGTATCATAAACGGTTTATGGS 1140
Db 1231 TAACTGGAAAGTAGAGGATTAATCAGATGACACCAACCGTATCATAAACGGTTTATGGS 1290
Qy 1141 GCAAGCGACCCGAAGAGAGAAATGCTAGCTACCATTTAGCTGTGTGGCGGCGCAAC 1200
Db 1291 GCAAGCGACCCGAAGAGAGAAATGCTAGCTATCATTTAGCCGTTGTTGCGGCGGAGC 1350
Qy 1201 AGATTGTAC 1209
Db 1351 AAATGGTTTC 1359
```

RESULT 4

US-09-940-235-6

; Sequence 6, Application US/09940235

; Publication No. US20030059921A1

; GENERAL INFORMATION:

```
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammar
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/09/940,235
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1327
; TYPE: DNA
; ORGANISM: Streptococcus equisimilis
; US-09-940-235-6

Query Match 76.5%; Score 1179; DB 3; Length 1327;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1182; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 TTTGTTTAACTTTAAGAAGGAGATATACCATGATAGCTGCTCTGTAATGCTACTAGATC 60
Db 51 TTTGTTTAACTTTAAGAAGGAGATATACCATGATAGCTGCTCTGTAATGCTACTAGATC 110
Qy 61 GTCCTTCTGTAAATAACAGCCAAATGCTGTTAGCGTGTCTGTACTGTTGAGGGGACGA 120
Db 111 GTCCTTCTGTAAATAACAGCCAAATGCTGTTAGCGTGTCTGTACTGTTGAGGGGACGA 170
Qy 121 ATCAAGACATTAAGCTTTAAATTTTGAATTCATCATCAACGACCTGCTCATGGAG 180
Db 171 ATCAAGACATTAAGCTTTAAATTTTGAATTCATCATCAACGACCTGCTCATGGAG 230
Qy 181 GAAAGACAGACGAGCGCTTAAGTCCAAATCAAAATCCATTTGCTACTGATAGTGGCGCGA 240
Db 231 GAAAGACAGACGAGCGCTTAAGTCCAAATCAAAATCCATTTGCTACTGATAGTGGCGCGA 290
Qy 241 TGTACATATAAATTTGAGAAAGCTGACTTACTAAAGGCTATTCAAGAACTATTGATCGCTA 300
Db 291 TGTACATATAAATTTGAGAAAGCTGACTTACTAAAGGCTATTCAAGAACTATTGATCGCTA 350
Qy 301 ACCTCCACAGTAACGACGACTACTTTGAGGTCAATTTGATTTTGCAGGATGCAACCAATTA 360
Db 351 ACCTCCACAGTAACGACGACTACTTTGAGGTCAATTTGATTTTGCAGGATGCAACCAATTA 410
Qy 361 CTGATCGAAACCGCAAGGCTCTACTTTGCTGACAAAGATGTTCCGTAACCTTGCCGACCC 420
Db 411 CTGATCGAAACCGCAAGGCTCTACTTTGCTGACAAAGATGTTCCGTAACCTTGCCGACCC 470
Qy 421 AACCTGTCCAAGAAATTTTGTGTAAGCGGACATGTGCGCGTTAGACCATATAAAGAAAC 480
Db 471 AACCTGTCCAAGAAATTTTGTGTAAGCGGACATGTGCGCGTTAGACCATATAAAGAAAC 530
Qy 481 CAATACAAACCAAGCGAAATCTGTTGATGTGGAATATCTGTACAGTTTACTCCCTTAA 540
Db 531 CAATACAAACCAAGCGAAATCTGTTGATGTGGAATATCTGTACAGTTTACTCCCTTAA 590
Qy 541 ACCCTGATGACGATTTACAGCAGGCTCTCAAGATACTAAGCTATTGAAACCACTAGCTA 600
Db 591 ACCCTGATGACGATTTACAGCAGGCTCTCAAGATACTAAGCTATTGAAACCACTAGCTA 650
Qy 601 TCGGTGACACCATCATCTCAAGAAATTAAGTCTCAAGCACAAGCAATTTTAAACAAA 660
Db 651 TCGGTGACACCATCATCTCAAGAAATTAAGTCTCAAGCACAAGCAATTTTAAACAAA 710
```

Qy	661	ACCACCCAGGCTATACGATTTATGAACGAGCTCCTCAATCGTCACTCATGACAATGACA	720
Db	711	ACCACCCAGGCTATACGATTTATGAACGAGCTCCTCAATCGTCACTCATGACAATGACA	770
Qy	721	TTTTCCGTCACGATTTTACCAATGGATCAAGAGTGTTCATTACCGTGTAAAAATCGGGAAC	780
Db	771	TTTTCCGTCACGATTTTACCAATGGATCAAGAGTGTTCATTACCGTGTAAAAATCGGGAAC	830
Qy	781	AAGCTTTATAGGATCAATAAAAAATCTGGTCTCGAATGAAGAAATAAACAACTGACCTGA	840
Db	831	AAGCTTTATAGGATCAATAAAAAATCTGGTCTCGAATGAAGAAATAAACAACTGACCTGA	890
Qy	841	TCTCTGAGAAATATTACGTCCTTAAAAAAGGGGAAAAGCCGTATGATCCCTTTGATCGCA	900
Db	891	TCTCTGAGAAATATTACGTCCTTAAAAAAGGGGAAAAGCCGTATGATCCCTTTGATCGCA	950
Qy	901	GTCACTTGAACCTGTTCCACCATCAATACGTTTGATGTCGATACCAACGAATTTGCTAAAAA	960
Db	951	GTCACTTGAACCTGTTCCACCATCAATACGTTTGATGTCGATACCAACGAATTTGCTAAAAA	1010
Qy	961	GTGAGCAGCTCTTAAACAGCTAGCGAAGTAACTTTAGACTTACAGAGATTTATACGATCCTC	1020
Db	1011	GTGAGCAGCTCTTAAACAGCTAGCGAAGTAACTTTAGACTTACAGAGATTTATACGATCCTC	1070
Qy	1021	GTGATAAGGCTAAACTACTCTACACAACTCTCGATGCTTTTGGTATTATGGACTATACCT	1080
Db	1071	GTGATAAGGCTAAACTACTCTACACAACTCTCGATGCTTTTGGTATTATGGACTATACCT	1130
Qy	1081	TAACTGGAAAAAGTAGAGGATTAATCAACGATGACCAACCGTATCATAAACCGTTTATATGG	1140
Db	1131	TAACTGGAAAAAGTAGAGGATTAATCAACGATGACCAACCGTATCATAAACCGTTTATATGG	1190
Qy	1141	GCAAGCGACCGAAGGAGAGAAATGCTAGCTACCATTTTACCTGGTGGT	1187
Db	1191	GCAAGCGACCGAAGGAGAGAAATGCTAGCTATCATTTAGCCTTATGAT	123

RESULT 5

```

US-09-940-235-5
; Sequence 5, Application US/09940235
; Publication No. US20030059921A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sabni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammara
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/09/940,235
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1377
; TYPE: DNA
; ORGANISM: Streptococcus equisimilis
US-09-940-235-5

```

```
Query Match          75.7%; Score 1166.2; DB 3; Length 1377;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1174; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
```

Qy	1	TTTTGTTTAACTTTTAAAGGAGGATATACCATGATAGCTGCTCTGGAATGCGTACTAGATC	60
Db	101	TTTTGTTTAACTTTTAAAGGAGGATATACCATGATGCTGGAACCTGAGTGGCTGCTAGACC	160
Qy	61	GTCCTTCTGTAAATAACAGGCCAATTGGTTGTTAGCGTTGCTGGTACTGTTGAGGGGACGA	120
Db	161	GTCCATCTGTCCACACAGCCCAATTGGTTGTTAGCGTTGCTGGTACTGTTGAGGGGACGA	220
Qy	121	ATCAAGACATTAGTCTTAAATTTTTTGAAATCGATCTAAACATACGACCTGCTCATGGAG	180
Db	221	ATCAAGACATTAGTCTTAAATTTTTTGAAATCGATCTAAACATACGACCTGCTCATGGAG	280
Qy	181	GAAGACAGACGAGCGCTTAAGTCCAAAAATCAAAACCAATTTCGTACTGATAGTGGCGCGA	240
Db	281	GAAGACAGACGAGCGCTTAAGTCCAAAAATCAAAACCAATTTCGTACTGATAGTGGCGCGA	340
Qy	241	TGTCACATAAATCTTGAGAAAGCTGACATTACTTAAAGGCTATTTCGAAGAACCAATTCATCGCTA	300
Db	341	TGTCACATAAATCTTGAGAAAGCTGACATTACTTAAAGGCTATTTCGAAGAACCAATTCATCGCTA	400
Qy	301	ACGTCCACAGTAACGACGACTACTTTTGAGGTCAATTGATTTTGAAGCGATGCAACCATTTA	360
Db	401	ACGTCCACAGTAACGACGACTACTTTTGAGGTCAATTGATTTTGAAGCGATGCAACCATTTA	460
Qy	361	CTGATCGAAACGGCAAGGCTCTACTTTGCTCGACAAAGATGGTTTCGGTAAACCTTCCGCAACC	420
Db	461	CTGATCGAAACGGCAAGGCTCTACTTTGCTCGACAAAGATGGTTTCGGTAAACCTTCCGCAACC	520
Qy	421	AACCTGTCCAGAAATTTTTCGTAAGCGGACATGTGCGCGTTAGACCATATAAAGAAAAAC	480
Db	521	AACCTGTCCAGAAATTTTTCGTAAGCGGACATGTGCGCGTTAGACCATATAAAGAAAAAC	580
Qy	481	CAATACAAAAACCAAGCGAAATCTGTTGATGTGGAATATACTGTACAGTTTACTTCCCTTAA	540
Db	581	CAATACAAAAACCAAGCGAAATCTGTTGATGTGGAATATACTGTACAGTTTACTTCCCTTAA	640
Qy	541	ACCTGATGAGATTTTCAGACCAAGTCTCAAGATACTTAAGTATTGAAACCACTAGCTA	600
Db	641	ACCTGATGAGATTTTCAGACCAAGTCTCAAGATACTTAAGTATTGAAACCACTAGCTA	700
Qy	601	TCGGTGCACCATCACAATCTCAAGAAATTAAGTCTCAAGCACAAGCATTTTAAACAAAA	660
Db	701	TCGGTGCACCATCACAATCTCAAGAAATTAAGTCTCAAGCACAAGCATTTTAAACAAAA	760
Qy	661	ACCACCAAGGCTATACGATTTATGAACGTCACTCCTCAATCGTCACTCATGACAAATGACA	720
Db	761	ACCACCAAGGCTATACGATTTATGAACGTCACTCCTCAATCGTCACTCATGACAAATGACA	820
Qy	721	TTTTCCGTACGATTTTACCAATGGATCAAGATTTTACTTACCGTGTATAAAATCGGGAAC	780
Db	821	TTTTCCGTACGATTTTACCAATGGATCAAGATTTTACTTACCGTGTATAAAATCGGGAAC	880
Qy	781	AAGCTTATAGGATCAATATAAAATCTGGTCTGAATGAAGAAATAAACAACATGACCTGA	840
Db	881	AAGCTTATAGGATCAATATAAAATCTGGTCTGAATGAAGAAATAAACAACATGACCTGA	940
Qy	841	TCTCTGAGAAATATTACGCTCTTAAAAAAGGGGAAAGCCGTATGATCCCTTTGATCGCA	900
Db	941	TCTCTGAGAAATATTACGCTCTTAAAAAAGGGGAAAGCCGTATGATCCCTTTGATCGCA	1000
Qy	901	GTCACTTTGAAACTGTTCCACCATCAAAATACGTTGATGTCGATACCAACGAATTCGTAAAAA	960
Db	1001	GTCACTTTGAAACTGTTCCACCATCAAAATACGTTGATGTCGATACCAACGAATTCGTAAAAA	1060
Qy	961	GTGAGCAGCTCTTAAACAGCTAGGCAACGTTAACTTTAGACTTTTACGAGTCTTC	1020
Db	1061	GTGAGCAGCTCTTAAACAGCTAGGCAACGTTAACTTTAGACTTTTACGAGTCTTC	1120
Qy	1021	GTGATAGGCTAACTACTCTCAACAATCTCGATGCTTTTGGTATTATGGACTATACCT	1080
Db	1121	GTGATAGGCTAACTACTCTCAACAATCTCGATGCTTTTGGTATTATGGACTATACCT	1180
Qy	1081	TAACTGAAAGTAGAGGATAATACGATGACACCAACCGTATCATAAACGTTTATATGG	1140

Db 1181 TAACTGAAAGTAGAGGATAATCAGATGACACCAACCGTATCATAAACGGTTATATGG 1240
 Qy 1141 GCAAGCGCCGAGGAGAGGAGTCTAGCTACCAATTTAGCTGGT 1187
 Db 1241 GCAAGCGCCGAGGAGAGGAGTCTAGCTATCATTTAGCCTATGAT 1287

RESULT 6

US-09-940-235-11
 ; Sequence 11, Application US/09940235
 ; Publication NO. US20030059921A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kumar, Rajesh
 ; APPLICANT: Sahni, Girish
 ; APPLICANT: Roy, Chait
 ; APPLICANT: Rajagopal, Kammara
 ; APPLICANT: Nihalani, Deepak
 ; APPLICANT: Sundaram, Vasudha
 ; APPLICANT: Yadav, Mahavir
 ; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
 ; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
 ; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
 ; TITLE OF INVENTION: PROTEIN
 ; FILE REFERENCE: 07064-009002
 ; CURRENT APPLICATION NUMBER: US/09/940,235
 ; PRIOR FILING DATE: 2002-04-09
 ; PRIOR FILING DATE: 1999-12-23
 ; PRIOR FILING DATE: 1998-12-24
 ; NUMBER OF SEQ ID NOS: 28
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 11
 ; LENGTH: 1782
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Hybrid cassette
 US-09-940-235-11

Query Match 73.6%; Score 1134.2; DB 3; Length 1782;
 Best Local Similarity 98.9%; Pred. No. 0;
 Matches 1142; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
 Qy 33 ATAGCTGGTCTGAAATGGCTACTAGATCGTCTTCTGTAATAACAGCCAAATTTGGTTT 92
 Db 538 ATTGCTGGACCTGAGTGGCTGCTAGACCGCTCCATCTGTCAACAACAGCCAAATTTGGTTT 597
 Qy 93 AGCGTTGCTGTTACTGTTGAGGGGACGAATCAAGACATAGTCTTAAATTTTGAATC 152
 Db 598 AGCGTTGCTGTTACTGTTGAGGGGACGAATCAAGACATAGTCTTAAATTTTGAATC 657
 Qy 153 GATCTAAATCAACGACTGCTCATGAGGAGAAACAGACAGCAAGCTTAACTCCAAATCA 212
 Db 658 GATCTAAATCAACGACTGCTCATGAGGAGAAACAGACAGCAAGCTTAACTCCAAATCA 717
 Qy 213 AAACCAATTTGCTACTGATAGTGGCGGATGTCAATAAATTTGAGAAAGCTGACTACTA 272
 Db 718 AAACCAATTTGCTACTGATAGTGGCGGATGTCAATAAATTTGAGAAAGCTGACTACTA 777
 Qy 273 AAGGCTATTCAAGAACAAATTTGATCGTCAAGTCCAGTAACGAGTACTTTGAGGTC 332
 Db 778 AAGGCTATTCAAGAACAAATTTGATCGTCAAGTCCAGTAACGAGTACTTTGAGGTC 837
 Qy 333 ATTGATTTTGAAGCGGATCAACATTTACTGATCGAAACGCAAGCTTACTTTGCTGAC 392
 Db 838 ATTGATTTTGAAGCGGATCAACATTTACTGATCGAAACGCAAGCTTACTTTGCTGAC 897
 Qy 393 AAAGATGGTTCGGTAACTTGGCGACCCCAACCTGTGCCAAGAAATTTTGTCTAAGCGGACAT 452
 Db 898 AAAGATGGTTCGGTAACTTGGCGACCCCAACCTGTGCCAAGAAATTTTGTCTAAGCGGACAT 957

RESULT 7

US-09-940-235-1
 ; Sequence 1, Application US/09940235
 ; Publication NO. US20030059921A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kumar, Rajesh
 ; APPLICANT: Sahni, Girish
 ; APPLICANT: Roy, Chait
 ; APPLICANT: Rajagopal, Kammara
 ; APPLICANT: Nihalani, Deepak
 ; APPLICANT: Sundaram, Vasudha
 ; APPLICANT: Yadav, Mahavir
 ; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
 ; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
 ; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
 ; TITLE OF INVENTION: PROTEIN
 ; FILE REFERENCE: 07064-009002
 ; CURRENT APPLICATION NUMBER: US/09/940,235
 ; CURRENT FILING DATE: 2002-04-09
 ; PRIOR APPLICATION NUMBER: 09/471,349

Qy 453 GTGCGCTTAGACCATATAAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 512
 Db 958 GTGCGCTTAGACCATATAAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 1017
 Qy 513 GAATATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTACAGACGATCTCAAA 572
 Db 1018 GAATATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTACAGACGATCTCAAA 1077
 Qy 573 GATACCTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTA 632
 Db 1078 GATACCTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTA 1137
 Qy 633 GCTCAAGCACAAGCATTTTAAACCAAAACCAACCCAGGCTATACGATTTTATGAACGTCAC 692
 Db 1138 GCTCAAGCACAAGCATTTTAAACCAAAACCAACCCAGGCTATACGATTTTATGAACGTCAC 1197
 Qy 693 TCCTCAATCGTCTACTCATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 752
 Db 1198 TCCTCAATCGTCTACTCATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 1257
 Qy 753 TTTTACTTACCGTGTAAAAATCGGAAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 812
 Db 1258 TTTTACTTACCGTGTAAAAATCGGAAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 1317
 Qy 813 AATGAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAAGGG 872
 Db 1318 AATGAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAAGGG 1377
 Qy 873 GAAAAAGCGGTATGATCCCTTTTGTATCGCAGTCACCTTGAACACTGTTTCAACATCAATACGTT 932
 Db 1378 GAAAAAGCGGTATGATCCCTTTTGTATCGCAGTCACCTTGAACACTGTTTCAACATCAATACGTT 1437
 Qy 933 GATGTCGATACCAACGAATTTGCTTAAAGTGAGAGCTCTTAAAGTGAGAGCTCTTAAAGTGAGAGCTTAA 992
 Db 1438 GATGTCGATACCAACGAATTTGCTTAAAGTGAGAGCTCTTAAAGTGAGAGCTCTTAAAGTGAGAGCTTAA 1497
 Qy 993 TTAGACTTCAGAGATTTATACGATCTCTCGTGATAAGGCTAACTTCTCAACAATCTC 1052
 Db 1498 TTAGACTTCAGAGATTTATACGATCTCTCGTGATAAGGCTAACTTCTCAACAATCTC 1557
 Qy 1053 GATGCTTTTGGTATTATGGAATATACCTTAACTTGAAGAGTAGAGGATAATCAGATGAC 1112
 Db 1558 GATGCTTTTGGTATTATGGAATATACCTTAACTTGAAGAGTAGAGGATAATCAGATGAC 1617
 Qy 1113 ACCAAGCTATCATACCGTTTATATGGCAAGGACCCGAGAGAGAGATGCTAGCTAC 1172
 Db 1618 ACCAAGCTATCATACCGTTTATATGGCAAGGACCCGAGAGAGAGATGCTAGCTAT 1677
 Qy 1173 CATTTAGCTGGTGGT 1187
 Db 1678 CATTTAGCTGGTGGT 1692

; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1245
; TYPE: DNA
; ORGANISM: Streptococcus equisimilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1242)
US-09-940-235-1

Query Match 73.5%; Score 1132.6; DB 3; Length 1245;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 1141; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 33 ATAGCTGCTCTGAATGGCTACTAGATCGTCTTCTGTAAATAACAGCCAAATTTGGTTGTT 92
Db |||||
Qy 93 AGCGTGTCTGCTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGGAAATC 152
Db |||||
Qy 61 AGCGTGTCTGCTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGGAAATC 120
Db |||||
Qy 153 GATCTAAACATACGACCTCTCTCATGAGGAGAAAGACAGAGCAAGGCTTAAGTCCAAATCA 212
Db |||||
Qy 121 GATCTAAACATACGACCTCTCTCATGAGGAGAAAGACAGAGCAAGGCTTAAGTCCAAATCA 180
Db |||||
Qy 213 AAACCATTTGCTACTGATAGTGGCGGATGTCAATAAACTTGAGAAAGCTGACTACTA 272
Db |||||
Qy 181 AAACCATTTGCTACTGATAGTGGCGGATGTCAATAAACTTGAGAAAGCTGACTACTA 240
Db |||||
Qy 273 AAGCTATTCAAGAAATGATCGCTAAGCTCAAGTAAACGACGACTACTTTGAGGTC 332
Db |||||
Qy 241 AAGCTATTCAAGAAATGATCGCTAAGCTCAAGTAAACGACGACTACTTTGAGGTC 300
Db |||||
Qy 333 ATTGATTTTGAAGGATGCAACCATTAAGTCTGAGAAAGGCAAGGCTACTTTGCTGAC 392
Db |||||
Qy 301 ATTGATTTTGAAGGATGCAACCATTAAGTCTGAGAAAGGCAAGGCTACTTTGCTGAC 360
Db |||||
Qy 393 AAAGATGTTTGGTAACTTGGCGACCAACCTGTCCAAAGAAATTTTGTCTAAGCGACAT 452
Db |||||
Qy 361 AAAGATGTTTGGTAACTTGGCGACCAACCTGTCCAAAGAAATTTTGTCTAAGCGACAT 420
Db |||||
Qy 453 GTGCGGTTAGACCATATAAAGAAACCAATAAAGAAACCAAGCGAAATCTGTGATGTG 512
Db |||||
Qy 421 GTGCGGTTAGACCATATAAAGAAACCAATAAAGAAACCAAGCGAAATCTGTGATGTG 480
Db |||||
Qy 513 GAATATCTGTACAGTTTACTTCCCTTAAACCTGTATGACGATTTGACAGGCTCTCAAA 572
Db |||||
Qy 481 GAATATCTGTACAGTTTACTTCCCTTAAACCTGTATGACGATTTGACAGGCTCTCAAA 540
Db |||||
Qy 573 GATCTAAGCTATTGAAACACTAGTATCGGTGACACCATCAATCTCAAGAAATTAATA 632
Db |||||
Qy 541 GATCTAAGCTATTGAAACACTAGTATCGGTGACACCATCAATCTCAAGAAATTAATA 600
Db |||||
Qy 633 GCTCAAGCACAAAGCATTTTAAACAAACCAACCAAGGCTATACGATTTTATGAACGTGAC 692
Db |||||
Qy 601 GCTCAAGCACAAAGCATTTTAAACAAACCAACCAAGGCTATACGATTTTATGAACGTGAC 660
Db |||||
Qy 693 TCCTCAATCTGTCATGACAAATGACATTTTCCGTAACGATTTTACCAATGGAATCAAGAG 752
Db |||||
Qy 661 TCCTCAATCTGTCATGACAAATGACATTTTCCGTAACGATTTTACCAATGGAATCAAGAG 720
Db |||||
Qy 753 TTTTCTTACCGTGTAAATAACCGGACCAAGCTTATAGATCAATAAATAATCTGCTCTG 812
Db |||||
Qy 721 TTTTCTTACCGTGTAAATAACCGGACCAAGCTTATAGATCAATAAATAATCTGCTCTG 780
Db |||||
Qy 813 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAAGTCTTAAATAAGGG 872
Db |||||
Qy 781 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAAGTCTTAAATAAGGG 840
Db |||||

RESULT 8

US-10-474-792-657
; Sequence 657, Application US/10474792
; Publication No. US20040236072A1
; GENERAL INFORMATION:
; APPLICANT: Olmsted, Stephen
; APPLICANT: Zagursky, Robert
; APPLICANT: Nickbarg, Elliot
; APPLICANT: Winter, Lourie
; TITLE OF INVENTION: SURFACE PROTEINS OF STREPTOCOCCUS PYOGENES
; FILE REFERENCE: AM 100399
; CURRENT APPLICATION NUMBER: US/10/474,792
; NUMBER OF SEQ ID NOS: 674
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 657
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Streptococcus pyogenes
US-10-474-792-657

Query Match 64.5%; Score 993.4; DB 8; Length 1323;
Best Local Similarity 91.3%; Pred. No. 6.4e-275;
Matches 1054; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

Qy 33 ATAGCTGGTCTCTGAATGGCTACTAGATCGTCTTCTGTAAATAACAGCCAAATTTGGTTGTT 92
Db |||||
Qy 79 ATTGCTGGGTATGGATGGCTTACCAGACCGTCCACCTATCAATACAGCCAGTGTAGTTGTT 138
Db |||||
Qy 93 AGCGTGTCTGCTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGGAAATC 152
Db |||||
Qy 139 AGTATGCGCGGTATCGTTGAAGGTACCGATAAAAAAGTTTATATAAATTTTGGAAATC 198
Db |||||
Qy 153 GATCTAATCATCAGACCTCTCATGAGGAGAAAGACAGAGCAAGGCTTAAGTCCAAATCA 212
Db |||||
Qy 199 GATCTAATCATCAACCTCTCAGGAGGAGAAAGACAGAGCAAGGCTTAAGTCCAAATCA 258
Db |||||
Qy 213 AAACCATTTTGTCTACTGATAGTGGCGGATGTCAATATACTTGAGAAAGCTGACTTACTA 272
Db |||||
Qy 259 AAACCATTTTGTCTACAGATAATGGGCAATGCCACATAAACTTGAAAGAGCTGACTTATTA 318
Db |||||
Qy 273 AAGCTATTCAAGAAACAAATTTGATCGCTTAAAGTCCAGTAAACGACGACTACTTTGAGGTC 332
Db |||||
Qy 319 AAAGCTATTCAAAAAACAGCTGATCGCTTAAAGTCCAGTAAACGACGCTACTTTGAGGTC 378
Db |||||
Qy 333 ATTGATTTTGAAGCGATGCAACCATTAAGTCTGAGAAAGGCAAGGCTTACTTTGCTGAC 392
Db |||||

```

Db 379 ATTGATTTTGAAGCGATGCAACCAATTAAGTATCGAAACGGCAAGGCTCTACTTTGCTGAC 438
QY 393 AAAGATGTTTGGTAACCTTGGCCGACCAACCTGTCCAAGAAATTTTGTCTAAGCGGACAT 452
Db 439 AAAGATGTTTGGTAACCTTGGCCGACCAACCTGTCCAAGAAATTTTGTCTAAGCGGACAT 498
QY 453 GTGCGGTTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTGTATGTG 512
Db 499 GTGCGGTTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTGTATGTG 558
QY 513 GAATATCTGTACGTTTACTCTCTTAAACCCCTGATGAGATTTAGACAGGCTCAAA 572
Db 559 GAATATCTGTACGTTTACTCTCTTAAACCCCTGATGAGATTTAGACAGGCTCAAA 618
QY 573 GATCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATTAATA 632
Db 619 GATCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATTAATA 678
QY 633 GCTCAAGCACAAAGCATTTTAAACAAAACCAACCCAGGCTATACGATTTTATGAACGTGAC 692
Db 679 GCTCAAGCACAAAGCATTTTAAACAAAACCAACCCAGGCTATACGATTTTATGAACGTGAC 738
QY 693 TCCTCAATCGTCACTCATGACAAATGACATTTTCGGTACGATTTTACCAATGGAATCAAGAG 752
Db 739 TCCTCAATCGTCACTCATGACAAATGACATTTTCGGTACGATTTTACCAATGGAATCAAGAG 798
QY 753 TTTTACTTACCGTGTAAAATCGGGAACAAGCTTATAGGATCAATAAAATCTGCTCTG 812
Db 799 TTTTACTTACCGTGTAAAATCGGGAACAAGCTTATAGGATCAATAAAATCTGCTCTG 858
QY 813 AATGAAGAAATAAACAACACTGACCTGTATCTGAGAAATATTAACGCTCTTAAAAAGGG 872
Db 859 AAGAAAAACGAACAACTGATCTGCTCTGAGAAATATTAACGCTCTTAAAAAGGG 918
QY 873 GAAAGCGGTATGATCCCTTTGATCGAGTCACTTGAAACTGTTCACCATCAATACGTT 932
Db 919 GAAAGCGGTATGATCCCTTTGATCGAGTCACTTGAAACTGTTCACCATCAATACGTT 978
QY 933 GATGTCGATACCAAGATTTGTAAGAGTGAAGCTCTTAAACAGCTAGCGAACGTAAC 992
Db 979 GATGTCGATACCAAGATTTGTAAGAGTGAAGCTCTTAAACAGCTAGCGAACGTAAC 1038
QY 993 TTAGACTTTAGAGATTTATACATCCTCGTGTATAGGCTTAACTACTCTACAAATCTC 1052
Db 1039 TTAGACTTTAGAGATTTATACATCCTCGTGTATAGGCTTAACTACTCTACAAATCTC 1098
QY 1053 GATGCTTTGGTATATGAGCTATACCTTTAACTGGAAAGTAGAGGATATCACGATGAC 1112
Db 1099 GATGCTTTGGTATATGAGCTATACCTTTAACTGGAAAGTAGAGGATATCACGATGAC 1158
QY 1113 ACCAACGCTATCATACCGTTTATATGGCAAGCGACCGAAGGAGAGATGCTAGCTAC 1172
Db 1159 AATAATCGTGTGTTACAGTTTATATGGCAAGCGCTTAAAGGGGCAAAAGGGTAGCTAT 1218
QY 1173 CATTTAGCTGGTGT 1187
Db 1219 CATTTAGCTTATGAT 1233

```

```

RESULT 9
US-10-210-235-3
; Sequence 3, Application US/09940235
; Publication No. US20030059921A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammaru
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION

```

```

; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/09/940,235
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 777
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(777)
US-09-940-235-3

Query Match 21.4%; Score 330; DB 3; Length 777;
Best Local Similarity 100.0%; Pred. No. 7.2e-84;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1209 CCCATAGCTGAGAAAGTGTGTTTGTATCATGCTGCTGGGACTTCCCTATGTGGTGGAGAAACG 1268
Db 448 CCCATAGCTGAGAAAGTGTGTTTGTATCATGCTGCTGGGACTTCCCTATGTGGTGGAGAAACG 507
QY 1269 TGGGAGAAAGCCCTTACCAAGCTGGATGATGGTATGATTTGTTCTTGGGAGAAAGGAGC 1328
Db 508 TGGGAGAAAGCCCTTACCAAGCTGGATGATGGTATGATTTGTTCTTGGGAGAAAGGAGC 567
QY 1329 GAGCGCATCTTTCGCACTTCTAGAAATAGATGCAACGATCAGGACACAAAGGACATCTAT 1388
Db 568 GAGCGCATCTTTCGCACTTCTAGAAATAGATGCAACGATCAGGACACAAAGGACATCTAT 627
QY 1389 AGAATTGGAGACACCTTGGAGCAAGAGGATAATCGAGAAACCTTCCAGTGTGATCTGC 1448
Db 628 AGAATTGGAGACACCTTGGAGCAAGAGGATAATCGAGAAACCTTCCAGTGTGATCTGC 687
QY 1449 ACAGGCAACCGCCGAGGAGAGTGGAAAGTGTGAGAGGACACCTTGTGCGAGACCAATCG 1508
Db 688 ACAGGCAACCGCCGAGGAGAGTGGAAAGTGTGAGAGGACACCTTGTGCGAGACCAATCG 747
QY 1509 AGCGATCTGGCCCCCTTCCCGATGTTTCT 1538
Db 748 AGCGATCTGGCCCCCTTCCCGATGTTTCT 777

```

```

RESULT 10
US-10-210-120-49
; Sequence 49, Application US/10210120
; Publication No. US20030175736A1
; GENERAL INFORMATION:
; APPLICANT: Chinnaiyan, Arul M.
; APPLICANT: Rubina, Mark A.
; APPLICANT: Sreekumar, Arun
; TITLE OF INVENTION: Expression Profile of Prostate Cancer
; FILE REFERENCE: UM-07221
; CURRENT APPLICATION NUMBER: US/10/210,120
; CURRENT FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: US 60/309,581
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: US 60/334,468
; PRIOR FILING DATE: 2001-11-15
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 49
; LENGTH: 2127
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-210-120-49

```


Db 61 GTCCCTTCGTAATAAACGCCAATTGGTTGTAGCGTTCGTGTA CTGTTGAGGGGACGA 120
Qy 121 ATCAAGACATTAGTCTTAAATTTTGAATCGATCTAAACATCAAGCACTGCTCATGGAG 180
Db 121 ATCAAGACATTAGTCTTAAATTTTGAATCGATCTAAACATCAAGCACTGCTCATGGAG 180
Qy 181 GAAAGACAGAGCAAGGCTTAAGTCCAAATTCAAAACCAATTTGCTACTGATAGTGGCGCA 240
Db 181 GAAAGACAGAGCAAGGCTTAAGTCCAAATTCAAAACCAATTTGCTACTGATAGTGGCGCA 240
Qy 241 TGTCACTAATACTTGGAAGCTGACTTAAAGCTATTCAAGAACTTCAAGAACTTCAATCGCTA 300
Db 241 TGTCACTAATACTTGGAAGCTGACTTAAAGCTATTCAAGAACTTCAAGAACTTCAATCGCTA 300
Qy 301 ACGTCCACAGTAACGACGACTACTTTGAGGTCAATTTGCAAGCGATGCAACCATTA 360
Db 301 ACGTCCACAGTAACGACGACTACTTTGAGGTCAATTTGCAAGCGATGCAACCATTA 360
Qy 361 CTGATCGAAACGGCAAGGCTACTTTGCTGACAAAGATGGTTGGTAACTTGGCCGACCC 420
Db 361 CTGATCGAAACGGCAAGGCTACTTTGCTGACAAAGATGGTTGGTAACTTGGCCGACCC 420
Qy 421 AACCTGTCCAGAAATTTTGTCTAAGCGGACATGTGCGGCTTAGACCATATAAGAAAAAC 480
Db 421 AACCTGTCCAGAAATTTTGTCTAAGCGGACATGTGCGGCTTAGACCATATAAGAAAAAC 480
Qy 481 CAATACAAAACCAAGCAATCTGTTGATGCGAATATCTGACAGTTTACTCCCTTAA 540
Db 481 CAATACAAAACCAAGCAATCTGTTGATGCGAATATCTGACAGTTTACTCCCTTAA 540
Qy 541 ACCCTGATGACGATTTTCAGACAGGCTCTCAAGATACTAAGCTATTGAAAACTAGCTA 600
Db 541 ACCCTGATGACGATTTTCAGACAGGCTCTCAAGATACTAAGCTATTGAAAACTAGCTA 600
Qy 601 TCGGTGACACATCATCTCAAGAACTTACTAGCTCAAGCAAAAGCAATTTTAAACAAA 660
Db 601 TCGGTGACACATCATCTCAAGAACTTACTAGCTCAAGCAAAAGCAATTTTAAACAAA 660
Qy 661 ACCACCGAGCTATACGATTTATGACGAGCTCTCAATCGTCTCAATCGTCTCAATGACA 720
Db 661 ACCACCGAGCTATACGATTTATGACGAGCTCTCAATCGTCTCAATCGTCTCAATGACA 720
Qy 721 TTTTCCGTACGATTTTACCAATGGATCAAGAGTTTACTTACCGTGTAAAAATCGGGAAC 780
Db 721 TTTTCCGTACGATTTTACCAATGGATCAAGAGTTTACTTACCGTGTAAAAATCGGGAAC 780
Qy 781 AAGCTTATAGGATCAATAAAAAATCTGGTCTGAATGAAGAAATAAACAACACTGACCTGA 840
Db 781 AAGCTTATAGGATCAATAAAAAATCTGGTCTGAATGAAGAAATAAACAACACTGACCTGA 840
Qy 841 TCTCTGAGAAATATTACGTCCTTAAAAAGGGGAAAGCCGTATGATCCCTTTGATCGCA 900
Db 841 TCTCTGAGAAATATTACGTCCTTAAAAAGGGGAAAGCCGTATGATCCCTTTGATCGCA 900
Qy 901 GTCACTTGAAACTGTTTCAACATCAATAGCTTGTGATGATACCAACGAATGCTAAAAA 960
Db 901 GTCACTTGAAACTGTTTCAACATCAATAGCTTGTGATGATACCAACGAATGCTAAAAA 960
Qy 961 GTGAGCAGCTCTTAAACAGCTAGCGAAGCTTAACTTAGACTTCAGAGATTTATACGATCCTC 1020
Db 961 GTGAGCAGCTCTTAAACAGCTAGCGAAGCTTAACTTAGACTTCAGAGATTTATACGATCCTC 1020
Qy 1021 GTGATAAGGCTTAACTTCTACAACTCTCGATGCTTTTGGTATTATGACTATACCT 1080
Db 1021 GTGATAAGGCTTAACTTCTACAACTCTCGATGCTTTTGGTATTATGACTATACCT 1080
Qy 1081 TAACTGGAAGAGTAGAGGATAATCAGATGACCAACCGTATCATACCGTTTATATGG 1140
Db 1081 TAACTGGAAGAGTAGAGGATAATCAGATGACCAACCGTATCATACCGTTTATATGG 1140
Qy 1141 GCAAGCGACCGGAAGAGAGATGCTAGCTACCATTTAGCTGTTGGTGGCGGCGCAAC 1200
Db 1141 GCAAGCGACCGGAAGAGAGATGCTAGCTACCATTTAGCTGTTGGTGGCGGCGCAAC 1200

Qy 1201 AGATTGTACCCATAGCTGAGAAGTGTGTTGATCATGCTGCTGGGACTTCCTATGTGTCG 1260
Db 1201 AGATTGTACCCATAGCTGAGAAGTGTGTTGATCATGCTGCTGGGACTTCCTATGTGTCG 1260
Qy 1261 GAGAAACGTGGGAGAGCCCTTACCAAGGCTGGATGATGGTAGATTGTACTTGCCTGGGAG 1320
Db 1261 GAGAAACGTGGGAGAGCCCTTACCAAGGCTGGATGATGGTAGATTGTACTTGCCTGGGAG 1320
Qy 1321 AAGGCGCGGACCCATCACTTGCACCTTCTAGAAATAGATGCAAGGATCAGGACACAGGA 1380
Db 1321 AAGGCGCGGACCCATCACTTGCACCTTCTAGAAATAGATGCAAGGATCAGGACACAGGA 1380
Qy 1381 CATCTTATAGAAATTGGAGACACCTGGAGCAAGAGGATTAATCGAGGAAACCTGCTCCAGT 1440
Db 1381 CATCTTATAGAAATTGGAGACACCTGGAGCAAGAGGATTAATCGAGGAAACCTGCTCCAGT 1440
Qy 1441 GCATCTGCACAGCAACGGCCGAGGAGAGTGGAAAGTGTGAGAGGCACACCTTGTGCGAGA 1500
Db 1441 GCATCTGCACAGCAACGGCCGAGGAGAGTGGAAAGTGTGAGAGGCACACCTTGTGCGAGA 1500
Qy 1501 CCACATCGAGCGGATCTGGCCCTTCCACCGATGTTGGTTAG 1541
Db 1501 CCACATCGAGCGGATCTGGCCCTTCCACCGATGTTGGTTAG 1541

RESULT 2
US-10-631-558-12
; Sequence 12, Application US/10631558
; Publication No. US20050260598A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammarra
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/10/631,558
; PRIOR FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US/09/940,235
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2096
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hybrid cassette
US-10-631-558-12

Query Match 97.1%; Score 1496.2; DB 7; Length 2096;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1501; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
Qy 33 ATAGCTGGTCTCTGAATGGCTACTAGATCGTCCTTCTGTAATAAATCAAGCAATTTGGTTGT 92
Db 588 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACCAATTTGGTTGT 647
Qy 93 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC 152
Db 648 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC 707

QY 153 GATCTAACATCAGACCTGCTCATGAGGAAAGACAGAGGCTTAAGTCCAAATCA 212
Db 708 GATCTAACATCAGACCTGCTCATGAGGAAAGACAGAGGCTTAAGTCCAAATCA 767
QY 213 AAACATTTGCTACTGATAGTGGCGGATGCTCATTAATTGAGAAAGCTGACTTA 272
Db 768 AAACATTTGCTACTGATAGTGGCGGATGCTCATTAATTGAGAAAGCTGACTTA 827
QY 273 AAGGCTATTCAAGAACATTTGATGCTTAAGCTCAAGTAAAGGCTTAAGTGAAGT 332
Db 828 AAGGCTATTCAAGAACATTTGATGCTTAAGCTCAAGTAAAGGCTTAAGTGAAGT 887
QY 333 ATTTGATTTTGCAAGCGATGCAACATTTGATGCTTAAGTAAAGGCTTAAGTGAAGT 392
Db 888 ATTTGATTTTGCAAGCGATGCAACATTTGATGCTTAAGTAAAGGCTTAAGTGAAGT 947
QY 393 AAAAGATGTTGGGTAACTTTGCGGACCCCACTGCTCAAGAAATTTTGTAGCGGACAT 452
Db 948 AAAAGATGTTGGGTAACTTTGCGGACCCCACTGCTCAAGAAATTTTGTAGCGGACAT 1007
QY 453 GTGGCGGTTAGACCATATTAAGAAACCAATACAAAGCGAAATCTGTTGATGTC 512
Db 1008 GTGGCGGTTAGACCATATTAAGAAACCAATACAAAGCGAAATCTGTTGATGTC 1067
QY 513 GAATATCTGTAAGATTTTACTCCCTTAAACCTGATGACGATTTGACAGGCTTCGAA 572
Db 1068 GAATATCTGTAAGATTTTACTCCCTTAAACCTGATGACGATTTGACAGGCTTCGAA 1127
QY 573 GATATCTAGCTTAAGTAAACATGATGCTGATGACCAATGATCAAGAAATTA 632
Db 1128 GATATCTAGCTTAAGTAAACATGATGCTGATGACCAATGATCAAGAAATTA 1187
QY 633 GGTCAAGCAAGGCTTTTAAACAAACCAAGCGGCTATGATTAAGAAAGCTGAC 692
Db 1188 GGTCAAGCAAGGCTTTTAAACAAACCAAGCGGCTATGATTAAGAAAGCTGAC 1247
QY 693 TCCTCAATGCTACTGATGACATGATGATTTTCCGTAAGATTTTCAATGATCAAG 752
Db 1248 TCCTCAATGCTACTGATGACATGATGATTTTCCGTAAGATTTTCAATGATCAAG 1307
QY 753 TTTACTTACCGTGTAAATGCGGAAACAAGCTTATGATCAATTAATGCTGCTG 812
Db 1308 TTTACTTACCGTGTAAATGCGGAAACAAGCTTATGATCAATTAATGCTGCTG 1367
QY 813 AATGAAGAAATTAACAAGCTGATGCTGGAAGATTTTACCTTTAAAGG 872
Db 1368 AATGAAGAAATTAACAAGCTGATGCTGGAAGATTTTACCTTTAAAGG 1427
QY 873 GAAAGCGGATGATCCCTTGTATGCGAGTCACTTGAAGCTGTCACATCAATGCTG 932
Db 1428 GAAAGCGGATGATCCCTTGTATGCGAGTCACTTGAAGCTGTCACATCAATGCTG 1487
QY 933 GATGTCATACCAAGAAATGCTTAAGAGCACTGTTAAACAGTACGAAAGTAA 992
Db 1488 GATGTCATACCAAGAAATGCTTAAGAGCACTGTTAAACAGTACGAAAGTAA 1547
QY 993 TTTAGCTTCAAGATTTATGATGCTGATGATGAGGCTTAAGTCAATCTC 1052
Db 1548 TTTAGCTTCAAGATTTATGATGCTGATGATGAGGCTTAAGTCAATCTC 1607
QY 1053 GATGCTTTTGTATTAAGATTAAGTAACTTAAAGTAAAGTAAAGTAAATCAAGTAC 1112
Db 1608 GATGCTTTTGTATTAAGATTAAGTAACTTAAAGTAAAGTAAAGTAAATCAAGTAC 1667
QY 1113 ACCAAGCGTATCATTAACGTTTATATGAGGCAAGCGGAGGAGGAGTCTAGTAC 1172
Db 1668 ACCAAGCGTATCATTAACGTTTATATGAGGCAAGCGGAGGAGGAGTCTAGTAC 1727
QY 1173 CATTTAGCTGATGAGGCAAGGAGGCAAGTGTATCCATAGCTGAGAAAGTCTTTGAT 1232
Db 1728 CATTTAGCTGATGAGGCAAGGAGGCAAGTGTATCCATAGCTGAGAAAGTCTTTGAT 1787
QY 1233 CATGCTGCTGGAGCTTCTATGATGATGAGAAAGCTGGAGAAAGCTTACCAAGGCTG 1292

Db 1788 CATGCTGCTGGAGCTTCTATGATGATGAGAAAGCTGGAGAAAGCTTACCAAGGCTG 1847
QY 1293 ATGATGATGATTTTACTGCTGAGGAGAAAGCGAGGAGGAGGAGTCTTGA 1352
Db 1848 ATGATGATGATTTTACTGCTGAGGAGAAAGCGAGGAGGAGGAGTCTTGA 1907
QY 1353 AATAGATGCAAGATGAGGAGCAAGAGGATCTTATAGATTTGAGAGCACTGAGCAAG 1412
Db 1908 AATAGATGCAAGATGAGGAGCAAGAGGATCTTATAGATTTGAGAGCACTGAGCAAG 1967
QY 1413 AAGATATGAGGAGAACTGCTCAGTGCATCTGCAAGGCAAGCGGAGGAGAGTGG 1472
Db 1968 AAGATATGAGGAGAACTGCTCAGTGCATCTGCAAGGCAAGCGGAGGAGAGTGG 2027
QY 1473 AAGTGTAGAGGACACCTCTGTCAGACCAATGAGGAGATCTGGCCCTTACCGAT 1532
Db 2028 AAGTGTAGAGGACACCTCTGTCAGACCAATGAGGAGATCTGGCCCTTACCGAT 2087
QY 1533 GTTCTTAG 1541
Db 2088 GTTCTTAG 2096

RESULT 3
US-10-631-558-10
; Sequence 10, Application US/10631558
; Publication No. US20050260598A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammarra
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vaesudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/10/631,558
; CURRENT FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US/09/940,235
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 1661
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURES:
; OTHER INFORMATION: Hybrid cassette
US-10-631-558-10

Query Match 76.9%; Score 1185; DB 7; Length 1661;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 1194; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 TTTGTTAACTTTAAGAGGATATATCAATGATAGCTGCTGAATGAGTACTAGATC 60
Db 151 TTTGTTAACTTTAAGAGGATATATCAATGATAGCTGCTGAATGAGTACTAGATC 210
QY 61 GTCTCTGTAAATTAACAGCAATGCTGTTAGCTGTTGCTGTTGAGGAGGAG 120
Db 211 GTCTCTGTAAATTAACAGCAATGCTGTTAGCTGTTGCTGTTGAGGAGGAG 270
QY 121 ATCAAGCATTAAGCTTAAATTTTGAATTCATCTAATCAAGCACTGCTCATGAG 180


```

Db      531 CAATACAAAACGAGGAAATCTGTTGATGTGAAATTAAGTACAGTTACTCTCTTAA 590
Qy      541 ACCCTGATGACGATTTTCAGACGAGTCTCAAGATTAAGTATTTGAAAACTAGCTA 600
Db      591 ACCCTGATGACGATTTTCAGACGAGTCTCAAGATTAAGTATTTGAAAACTAGCTA 650
Qy      601 TCGGTGACACCATCATCTCAAGAAATTAAGTCAAGCAAAAGATTTTAAACAAA 660
Db      651 TCGGTGACACCATCATCTCAAGAAATTAAGTCAAGCAAAAGATTTTAAACAAA 710
Qy      661 ACCACCGAGCTATTAAGTATTAAGACGTGCTCCCAATCTGCACTGACATGACATA 720
Db      711 ACCACCGAGCTATTAAGTATTAAGACGTGCTCCCAATCTGCACTGACATGACATA 770
Qy      721 TTTTCGTAAGATTTTACCAATGATCAAGATTTTACCTGCTTAAATCGGAGAC 780
Db      771 TTTTCGTAAGATTTTACCAATGATCAAGATTTTACCTGCTTAAATCGGAGAC 830
Qy      781 AAGCTTATAGATCAATTAATAAAATCTGCTGAAATGAATAAACAACACTGACCTGA 840
Db      831 AAGCTTATAGATCAATTAATAAAATCTGCTGAAATGAATAAACAACACTGACCTGA 890
Qy      841 TCTCTGAGAAATTTACGCTCTTAAAAAAGGGGAAAAAGCCGTATGATCCCTTGATCGCA 900
Db      891 TCTCTGAGAAATTTACGCTCTTAAAAAAGGGGAAAAAGCCGTATGATCCCTTGATCGCA 950
Qy      901 GTCACTTGAACCTGTTACCATCAATTAAGTATGATCAACCAAGATTTGCTTAAAA 960
Db      951 GTCACTTGAACCTGTTACCATCAATTAAGTATGATCAACCAAGATTTGCTTAAAA 1010
Qy      961 GTGAGAGCTCTTAAAGTAGAGCAAGTAACTTAGACTTCAGAGATTTTACGATCTC 1020
Db      1011 GTGAGAGCTCTTAAAGTAGAGCAAGTAACTTAGACTTCAGAGATTTTACGATCTC 1070
Qy      1021 GTGATTAAGCTTAACTCTCTCAACAATCTCGATGCTTTGGTATTAAGACTTAACT 1080
Db      1071 GTGATTAAGCTTAACTCTCTCAACAATCTCGATGCTTTGGTATTAAGACTTAACT 1130
Qy      1081 TAACTGGAAGAGTAGATTAATCAAGATGACCAACCGTATCAATACCGTTATATG 1140
Db      1131 TAACTGGAAGAGTAGATTAATCAAGATGACCAACCGTATCAATACCGTTATATG 1190
Qy      1141 GCAACGACCCGAGAGAGAAATGTAGCTACCATTTAGCTGGT 1187
Db      1191 GCAACGACCCGAGAGAGAAATGTAGCTACCATTTAGCTATGAT 1237

```

```

RESULT 5
US-10-631-558-5
; Sequence 5, Application US/10631558
; Publication No. US20050260598A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammaru
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEIN POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT FILING DATE: 2003-07-31
; PRIOR FILING DATE: 2002-04-09
; PRIOR FILING DATE: 1999-12-23
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28

```

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1377
; TYPE: DNA
; ORGANISM: Streptococcus equisimilis
US-10-631-558-5

```

```

Query Match      75.7%; Score 1166.2; DB 7; Length 1377;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1174; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

```

```

Qy      1 TTTGTTTAACTTTTAAAGAGATTAATACCATGATAGCTGCTGAAATGCTACTAGATC 60
Db      101 TTTGTTTAACTTTTAAAGAGATTAATACCATGATAGCTGCTGAAATGCTACTAGATC 160
Qy      61 GTCTCTGTAATTAACAGCCAAATGCTGTTGAGCTGCTGCTGCTGCTGCTGCTGCTG 120
Db      161 GTCTCTGTAATTAACAGCCAAATGCTGTTGAGCTGCTGCTGCTGCTGCTGCTGCTG 220
Qy      121 ATCAAGACATTAATTTTAAATTTTGAATCGATCTAATCAATCAATCAATCAATCAAT 180
Db      221 ATCAAGACATTAATTTTAAATTTTGAATCGATCTAATCAATCAATCAATCAATCAAT 280
Qy      181 GAAAGACAGAGCAAGCTTAAATGTCCTAAATCAAAACCAATTTGCTACTGATAGTGGCGGA 240
Db      281 GAAAGACAGAGCAAGCTTAAATGTCCTAAATCAAAACCAATTTGCTACTGATAGTGGCGGA 340
Qy      241 TGTCAATTAATCTTGAAGAGCTGATCTTAAAGCTATTAAGCTATTAAGCTATTAAG 300
Db      341 TGTCAATTAATCTTGAAGAGCTGATCTTAAAGCTATTAAGCTATTAAGCTATTAAG 400
Qy      301 AAGTCACAGTAAGAGAGCACTTAAATTTTGAATCGATCTAATCAATCAATCAATCAAT 360
Db      401 AAGTCACAGTAAGAGAGCACTTAAATTTTGAATCGATCTAATCAATCAATCAATCAAT 460
Qy      361 CTGATCGAAAACGAGCAAGCTCTAATTTGCTGCAAAAGAGTGGTGGTAACTTGGCGGACC 420
Db      461 CTGATCGAAAACGAGCAAGCTCTAATTTGCTGCAAAAGAGTGGTGGTAACTTGGCGGACC 520
Qy      421 AACTGTCCAAAGAAATTTTGTAAAGCGGACATGTGGCGTTAAACATTAAGAAAAAC 480
Db      521 AACTGTCCAAAGAAATTTTGTAAAGCGGACATGTGGCGTTAAACATTAAGAAAAAC 580
Qy      481 CAATTAAGAAACCAAGCGGAAATCTGTGATGAGAAATATATCTGATCACTTATCTCCCTTAA 540
Db      581 CAATTAAGAAACCAAGCGGAAATCTGTGATGAGAAATATATCTGATCACTTATCTCCCTTAA 640
Qy      541 ACCGTATGAGATTTTCAAGACGAGTCTCAAGATTAAGCTATTAAGCTATTAAGCTA 600
Db      641 ACCGTATGAGATTTTCAAGACGAGTCTCAAGATTAAGCTATTAAGCTATTAAGCTA 700
Qy      601 TCGGTGACACCATCATCTCAAGAAATTAAGTCAAGCAAAAGATTTTAAACAAA 660
Db      701 TCGGTGACACCATCATCTCAAGAAATTAAGTCAAGCAAAAGATTTTAAACAAA 760
Qy      661 ACCACCGAGCTATTAAGTATTAAGACGTGCTCCCAATCTGCACTGACATGACATA 720
Db      761 ACCACCGAGCTATTAAGTATTAAGACGTGCTCCCAATCTGCACTGACATGACATA 820
Qy      721 TTTTCGTAAGATTTTACCAATGATTAAGACGTGCTCCCAATCTGCACTGACATGACATA 780
Db      821 TTTTCGTAAGATTTTACCAATGATTAAGACGTGCTCCCAATCTGCACTGACATGACATA 880
Qy      781 AAGCTTATAGATCAATTAATAAAATCTGCTGAAATGAATAAACAACACTGACCTGA 840
Db      881 AAGCTTATAGATCAATTAATAAAATCTGCTGAAATGAATAAACAACACTGACCTGA 940
Qy      841 TCTCTGAGAAATTTACGCTCTTAAAAAAGGGGAAAAAGCCGTATGATCCCTTGATCGCA 900
Db      941 TCTCTGAGAAATTTACGCTCTTAAAAAAGGGGAAAAAGCCGTATGATCCCTTGATCGCA 1000
Qy      901 GTCACTTGAACCTGTTACCATCAATTAAGTATGATCAACCAAGATTTGCTTAAAA 960

```

Db 1001 GTCACTTGAAACTGTTCACCATCAAAATAGCTTGAATGTCGATACCAACGAATTGCTAAAAA 1060
Qy 961 GTGAGCAGCTCTTAACAGCTACGGAACGTAACCTTAGACTTCAGAGATTTATACGATCCCTC 1020
Db 1061 GTGAGCAGCTCTTAACAGCTACGGAACGTAACCTTAGACTTCAGAGATTTATACGATCCCTC 1120
Qy 1021 GTGATAAGGCTAAACTACTCTCTAACAACATCTCGATCTTTGGTATTTATGGACTATACCT 1080
Db 1121 GTGATAAGGCTAAACTACTCTCTAACAACATCTCGATCTTTGGTATTTATGGACTATACCT 1180
Qy 1081 TAACTGGAAGTAGAGGATATACGATGACCAACCGTATCATACACCGTTTATATGG 1140
Db 1181 TAACTGGAAGTAGAGGATATACGATGACCAACCGTATCATACACCGTTTATATGG 1240
Qy 1141 GCAAGCGACCCGAGGAGATGCTAGCTACCATTTAGCTGGTGGT 1187
Db 1241 GCAAGCGACCCGAGGAGATGCTAGCTACCATTTAGCTGGTGGT 1287

RESULT 6

US-10-631-558-11

; Sequence 11, Application US/10631558
; Publication No. US20050260598A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammarra
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/10/631,558
; CURRENT FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US/09/940,235
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 1782
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hybrid cassette

US-10-631-558-11

Query Match 73.6%; Score 1134.2; DB 7; Length 1782;
Best Local Similarity 98.9%; Pred. No. 0; Mismatches 13; Indels 0; Gaps 0;
Matches 1142; Conservative 0;

Qy 33 ATAGCTGCTCGTGAATGGCTACTAGATCGTCTCTCTGTAATAACAGCAATTTGGTTGT 92
Db 538 ATTGCTGACCTGAGTGGCTGTAGACCTCCATCTGTCAACACGACCAATTTGGTTGT 597
Qy 93 AGCGTTGCTGTTGTTGAGGGGACGAATCAAGACATTTAGTCTTAAATTTTTTGAATC 152
Db 598 AGCGTTGCTGTTGTTGAGGGGACGAATCAAGACATTTAGTCTTAAATTTTTTGAATC 657
Qy 153 GATCTAATCACCAGACCTGCTCATGAGGAGAAAGACAGCAAGGCTTAAGTCCAAATCA 212
Db 658 GATCTAATCACCAGACCTGCTCATGAGGAGAAAGACAGCAAGGCTTAAGTCCAAATCA 717
Qy 213 AAACCATTTGCTACTGATAGTGGCGGATGTCACATAAACTTGAGAAAGCTGACTTACTA 272
Db 718 AAACCATTTGCTACTGATAGTGGCGGATGTCACATAAACTTGAGAAAGCTGACTTACTA 777

RESULT 7

US-10-631-558-1
; Sequence 1, Application US/10631558
; Publication No. US20050260598A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish

Qy 273 AAGGCTATTCAAGAAACAATTTGATCGCTAACAGTCCACAGTAACGACGACTACTTTGAGGTC 332
Db 778 AAGGCTATTCAAGAAACAATTTGATCGCTAACAGTCCACAGTAACGACGACTACTTTGAGGTC 837
Qy 333 ATTGATTTTTCAAGCGATGCAACCAATTAATGATCGAAACGGCAAGGCTACTTTGCTGAC 392
Db 838 ATTGATTTTTCAAGCGATGCAACCAATTAATGATCGAAACGGCAAGGCTACTTTGCTGAC 897
Qy 393 AAGATGGTTCGTAACCTTGGCGACCCAACTTCTGATCGAAACGGCAAGGCTACTTTGCTGAC 452
Db 898 AAGATGGTTCGTAACCTTGGCGACCCAACTTCTGATCGAAACGGCAAGGCTACTTTGCTGAC 957
Qy 453 GTGCGGTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 512
Db 958 GTGCGGTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 1017
Qy 513 GAATATACTGTACAGTTTACTCCCTTTAAACCCCTGATGACGATTTTACAGCAGGCTCTCAAA 572
Db 1018 GAATATACTGTACAGTTTACTCCCTTTAAACCCCTGATGACGATTTTACAGCAGGCTCTCAAA 1077
Qy 573 GATACTAAGCTATGAAAAACATAGCTATCGGTGACACCATCATCTCAAGAAATTACTA 632
Db 1078 GATACTAAGCTATGAAAAACATAGCTATCGGTGACACCATCATCTCAAGAAATTACTA 1137
Qy 633 GCTCAAGCACAAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTTATGAACGTCGAC 692
Db 1138 GCTCAAGCACAAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTTATGAACGTCGAC 1197
Qy 693 TCCTCAATCGTCACTCATGACATGACATTTTCCGTCGATCGATTTTACCAATGGATCAAGAG 752
Db 1198 TCCTCAATCGTCACTCATGACATGACATTTTCCGTCGATCGATTTTACCAATGGATCAAGAG 1257
Qy 753 TTTACTTACCGTGTAAAAAATCGGGAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 812
Db 1258 TTTACTTACCGTGTAAAAAATCGGGAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 1317
Qy 813 AATGAAGAAATAACCAACTGACCTGATCTCTGAGAAATATTAACGTCCTTAAAAAGGG 872
Db 1318 AATGAAGAAATAACCAACTGACCTGATCTCTGAGAAATATTAACGTCCTTAAAAAGGG 1377
Qy 873 GAAAAGCGTATGATCCCTTTGATCGCAGTCACCTTGAAACCTGTTCAACATCAATAGCTT 932
Db 1378 GAAAAGCGTATGATCCCTTTGATCGCAGTCACCTTGAAACCTGTTCAACATCAATAGCTT 1437
Qy 933 GATGTCGATACCAACGAATTTGCTTAAAGAGTGACGAGCTCTTAAACAGCTAGCGAACGTAAC 992
Db 1438 GATGTCGATACCAACGAATTTGCTTAAAGAGTGACGAGCTCTTAAACAGCTAGCGAACGTAAC 1497
Qy 993 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTCAACAATCTC 1052
Db 1498 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTCAACAATCTC 1557
Qy 1053 GATGCTTTTGGTATTATGGAATATACCTTAACTGGAAAGTAGAGGATAATCAGATGAC 1112
Db 1558 GATGCTTTTGGTATTATGGAATATACCTTAACTGGAAAGTAGAGGATAATCAGATGAC 1617
Qy 1113 ACCAAGCGTATCAATACCGTTTATATGGCAAGCGACCCGAGGAGAGAAATGCTAGTAC 1172
Db 1618 ACCAAGCGTATCAATACCGTTTATATGGCAAGCGACCCGAGGAGAGAAATGCTAGTAC 1677
Qy 1173 CATTTAGCTGGTGGT 1187
Db 1678 CATTTAGCTGGTGGT 1692

```

1  APPLICANT: Roy, Chait
2  APPLICANT: Rajagopal, Kammara
3  APPLICANT: Nihalani, Deepak
4  APPLICANT: Sundaram, Vasudha
5  APPLICANT: Yadav, Mahavir
6  TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
7  TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
8  TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAIDIDS
9  TITLE OF INVENTION: PROTEIN
10 FILE REFERENCE: 07064-009002
11 CURRENT APPLICATION NUMBER: US/10/631,558
12 CURRENT FILING DATE: 2003-07-31
13 PRIOR APPLICATION NUMBER: US/09/940,235
14 PRIOR FILING DATE: 2002-04-09
15 PRIOR APPLICATION NUMBER: 09/471,349
16 PRIOR FILING DATE: 1999-12-23
17 PRIOR APPLICATION NUMBER: IN 3825/DEL/98
18 PRIOR FILING DATE: 1998-12-24
19 NUMBER OF SEQ ID NOS: 28
20 SOFTWARE: FastSeq for Windows Version 4.0
21 SEQ ID NO 1
22 LENGTH: 1245
23 TYPE: DNA
24 ORGANISM: Streptococcus equisimilis
25 FEATURE:
26 NAME/KEY: CDS
27 LOCATION: (1)...(1242)
28 US-10-631-558-1

```

Query Match	Similarity	78.5%	Score 1132.6	DB 7	Length 1245	
Best Local	Similarity	93.8%	Pred. No. 0			
Matches 1141	Conservative	0	Mismatches 14	Indels 0	Gaps 0	
QY	33	ATAGTGTCTCTGAATGCGTACTGATGCGTCTCTGTATAATATACAGCCAAATGGTGT	92			
Db	1	ATTGTGTGACCTGAGTGGCTGTGACCGTCCATCTGTCTCAACAAACAGCCAAATTAGTGT	60			
QY	93	AGCGTGTCTGTACTGTGTAGGGGAGCGAATCAAGCATTAAGTCTTAATTTTTGAATC	152			
Db	61	AGCGTGTCTGTACTGTGTAGGGGAGCGAATCAAGCATTAAGTCTTAATTTTTGAATC	120			
QY	153	GATCTAACATCAACGACTGTCTCATGAGAGAAACAGAGCAAGCTTAAGTCCAAATCA	212			
Db	121	GATCTAACATCAACGACTGTCTCATGAGAGAAACAGAGCAAGCTTAAGTCCAAATCA	180			
QY	213	AAACCATTTTGTCTACTGATATGAGCGCGATGTCTACATTAACCTTGAGAAAGCTGACTTA	272			
Db	181	AAACCATTTTGTCTACTGATATGAGCGCGATGTCTACATTAACCTTGAGAAAGCTGACTTA	240			
QY	273	AAGGCTATTCAAGAACAAATTGATGCGTAAAGTCCACAGTAAACGAGACTACTTTGAGGTC	332			
Db	241	AAGGCTATTCAAGAACAAATTGATGCGTAAAGTCCACAGTAAACGAGACTACTTTGAGGTC	300			
QY	333	ATTGATTTTGGCAAGGATGCAACCATTACTGATCGAAAACGGCAAGTCTACTTTGCTGAC	392			
Db	301	ATTGATTTTGGCAAGGATGCAACCATTACTGATCGAAAACGGCAAGTCTACTTTGCTGAC	360			
QY	393	AAAGATGTTTGGGTAACTTGTCCGACCCAACTGTCTCAAGAAATTTTGTCTAAGCGGACAT	452			
Db	361	AAAGATGTTTGGGTAACTTGTCCGACCCAACTGTCTCAAGAAATTTTGTCTAAGCGGACAT	420			
QY	453	GTGCGCGTTAGACCATATATAAGAAAAACCAATTAACAAACCAAGCAAAATCTGTGATGTG	512			
Db	421	GTGCGCGTTAGACCATATATAAGAAAAACCAATTAACAAACCAAGCAAAATCTGTGATGTG	480			
QY	513	GAATATACGTGACAGTTTACTTCCCTTAAACCTTGATGACGATTTCAACCAAGTCTCAAA	572			
Db	481	GAATATACGTGACAGTTTACTTCCCTTAAACCTTGATGACGATTTCAACCAAGTCTCAAA	540			
QY	573	GATGCTAAGCATTTGAAAAACCTGAGCTATGGGTGACACATCAACACTTCAAGAAATTA	632			
Db	541	GATGCTAAGCATTTGAAAAACCTGAGCTATGGGTGACACATCAACACTTCAAGAAATTA	600			

OY	633	GCTAAGACAAGAATTTTAAACAAAACCAACCAGGCTATACGATTTATGAACGTCAC	692
Db	601	GCTCAGACACAAGAGATTTTAAACAAAACCAACCAGGCTATACGATTTATGAACGTCAC	660
OY	693	TCCTCAATCGTCACCTCAGACATGACATTTTCCGTACGATTTTACCAATGATCAAGAG	752
Db	661	TCCTCAATCGTCACATGACATGACATTTTCCGTACGATTTTACCAATGATCAAGAG	720
OY	753	TTTACTTAACCGTGTAAAAATCGGGAACAGCTTATAGATCAATAAAAATCTGCTCG	812
Db	721	TTTACTTAACCGTGTAAAAATCGGGAACAGCTTATAGATCAATAAAAATCTGCTCG	780
OY	813	AATAGAGAAATPAAACAACACTGACCTGATCTCTGAGAAATTTTACGTCCTTAAAAAGGG	872
Db	781	AATAGAGAAATPAAACAACACTGACCTGATCTCTGAGAAATTTTACGTCCTTAAAAAGGG	840
OY	873	GAAAAGCGATGATCCCTTGTGATCGACAGTCACCTTGAAACCTGACCATCAATACGTT	932
Db	841	GAAAAGCGATGATCCCTTGTGATCGACAGTCACCTTGAAACCTGACCATCAATACGTT	900
OY	933	GATGTCGATACCAACGAATGCTTAAAAAGTGAAGACCTCTTAACAAGTACGGAACGTAC	992
Db	901	GATGTCGATACCAACGAATGCTTAAAAAGTGAAGACCTCTTAACAAGTACGGAACGTAC	960
OY	993	TTAGACTTCAGAGATTTATACGATCCGATGATGAAGGCTAAACCTACCTCAACAATCTC	1052
Db	961	TTAGACTTCAGAGATTTATACGATCCGATGATGAAGGCTAAACCTACCTCAACAATCTC	1020
OY	1053	GATGCTTTGGTATTTATGACCTATACCTTAACTGGAAGAGTGAAGAGATTAATCAGATGAC	1112
Db	1021	GATGCTTTGGTATTTATGACCTATACCTTAACTGGAAGAGTGAAGAGATTAATCAGATGAC	1080
OY	1113	ACCAACCGTATCATTAACCGTTTATATGGGCAAGGACCCGAAAGAGAGATGCTAGCTAC	1172
Db	1081	ACCAACCGTATCATTAACCGTTTATATGGGCAAGGACCCGAAAGAGAGATGCTAGCTAT	1140
OY	1173	CATTAGCTGGTGGT 1187	
Db	1141	CATTAGCTATGAT 1155	

```

RESULT 8
US-10-631-558-3
? Sequence 3, Application US/10631558
? Publication No. US20050260598A1
? GENERAL INFORMATION:
? APPLICANT: Kumar, Rajesh
? APPLICANT: Sahni, Girish
? APPLICANT: Roy, Chait
? APPLICANT: Rajagopal, Kammarra
? APPLICANT: Nihalani, Deepak
? APPLICANT: Sundaram, Vasudha
? APPLICANT: Yadav, Mahavir
? TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
? TITLE OF INVENTION: PROTEINS POSSESSING ALTEPED PLASMINOGEN ACTIVATION
? TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
? TITLE OF INVENTION: PROTEIN
? FILE REFERENCE: 07064-009002
? CURRENT APPLICATION NUMBER: US/10/631,558
? CURRENT FILING DATE: 2003-07-31
? PRIOR APPLICATION NUMBER: US/09/940,235
? PRIOR FILING DATE: 2002-04-09
? PRIOR APPLICATION NUMBER: 09/471,349
? PRIOR FILING DATE: 1999-12-23
? PRIOR APPLICATION NUMBER: IN 3825/DEL/98
? PRIOR FILING DATE: 1998-12-24
? NUMBER OF SEQ ID NOS: 28
? SOFTWARE: FastSeq for windows Version 4.0
? SEQ ID NO 3
? LENGTH: 777
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:

```

; NAME/KEY: CDS
; LOCATION: (1) ... (777)
US-10-631-558-3

Query Match 21.4%; Score 330; DB 7; Length 777;
Best Local Similarity 100.0%; Pred. No. 4.4e-95;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1209 CCCATAGCTGAGAAGTGTTCATCATGCTGCTGGGACTTCCTATGTGTGTCGGAGAAACG 1268
DB 448 CCCATAGCTGAGAAGTGTTCATCATGCTGCTGGGACTTCCTATGTGTGTCGGAGAAACG 507
QY 1269 TGGGAGAAGCCCTACCAAGCTGGGATGATGGTAGATTGCTCCCTGGGAGAGGAGCAGC 1328
DB 508 TGGGAGAAGCCCTACCAAGCTGGGATGATGGTAGATTGCTCCCTGGGAGAGGAGCAGC 567
QY 1329 GGAGCGATCATTGTCACCTTCTAGAAATAGATGCAACGATCAGGACACAAGGACATCCTAT 1388
DB 568 GGAGCGATCATTGTCACCTTCTAGAAATAGATGCAACGATCAGGACACAAGGACATCCTAT 627
QY 1389 AGAATTGGAGACACCTGGAGGAGAGTGAAGTGTGAGAGGACACCTCTGTGTCAGGACATCG 1448
DB 628 AGAATTGGAGACACCTGGAGGAGAGTGAAGTGTGAGAGGACACCTCTGTGTCAGGACATCG 687
QY 1449 ACAGGCAACGGCCGAGGAGAGTGAAGTGTGAGAGGACACCTCTGTGTCAGGACATCG 1508
DB 688 ACAGGCAACGGCCGAGGAGAGTGAAGTGTGAGAGGACACCTCTGTGTCAGGACATCG 747
QY 1509 AGCGGATCGGCCCTTCACCGATGTTGCT 1538
DB 748 AGCGGATCGGCCCTTCACCGATGTTGCT 777

RESULT 9

US-10-995-561-114
; Sequence 114, Application US/10995561
; Publication No. US20050272054A1

; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 114
; LENGTH: 2443
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-114

Query Match 21.4%; Score 330; DB 7; Length 2443;
Best Local Similarity 100.0%; Pred. No. 9.1e-95;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1209 CCCATAGCTGAGAAGTGTTCATCATGCTGCTGGGACTTCCTATGTGTGTCGGAGAAACG 1268
DB 913 CCCATAGCTGAGAAGTGTTCATCATGCTGCTGGGACTTCCTATGTGTGTCGGAGAAACG 972
QY 1269 TGGGAGAAGCCCTACCAAGCTGGGATGATGGTAGATTGCTCCCTGGGAGAGGAGCAGC 1328
DB 973 TGGGAGAAGCCCTACCAAGCTGGGATGATGGTAGATTGCTCCCTGGGAGAGGAGCAGC 1032
QY 1329 GGAGCGATCATTGTCACCTTCTAGAAATAGATGCAACGATCAGGACACAAGGACATCCTAT 1388
DB 1033 GGAGCGATCATTGTCACCTTCTAGAAATAGATGCAACGATCAGGACACAAGGACATCCTAT 1092
QY 1389 AGAATTGGAGACACCTGGAGGAGAGGATATCGAGGAAACCTGTCTCCAGTGCATCTGC 1448
DB 1093 AGAATTGGAGACACCTGGAGGAGAGGATATCGAGGAAACCTGTCTCCAGTGCATCTGC 1152

QY 1449 ACAGGCAACGGCCGAGGAGAGTGAAGTGTGAGAGGACACCTCTGTGTCAGACCATCG 1508
DB 1153 ACAGGCAACGGCCGAGGAGAGTGAAGTGTGAGAGGACACCTCTGTGTCAGACCATCG 1212
QY 1509 AGCGGATCTGGCCCCCTTCACCGATGTTGCT 1538
DB 1213 AGCGGATCTGGCCCCCTTCACCGATGTTGCT 1242

RESULT 10

US-10-995-561-105
; Sequence 105, Application US/10995561
; Publication No. US20050272054A1

; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 105
; LENGTH: 2488
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-105

Query Match 21.4%; Score 330; DB 7; Length 2488;

Best Local Similarity 100.0%; Pred. No. 9.2e-95;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1209 CCCATAGCTGAGAAGTGTTCATCATGCTGCTGGGACTTCCTATGTGTGTCGGAGAAACG 1268
DB 913 CCCATAGCTGAGAAGTGTTCATCATGCTGCTGGGACTTCCTATGTGTGTCGGAGAAACG 972
QY 1269 TGGGAGAAGCCCTACCAAGCTGGGATGATGGTAGATTGCTCCCTGGGAGAGGAGCAGC 1328
DB 973 TGGGAGAAGCCCTACCAAGCTGGGATGATGGTAGATTGCTCCCTGGGAGAGGAGCAGC 1032
QY 1329 GGAGCGATCATTGTCACCTTCTAGAAATAGATGCAACGATCAGGACACAAGGACATCCTAT 1388
DB 1033 GGAGCGATCATTGTCACCTTCTAGAAATAGATGCAACGATCAGGACACAAGGACATCCTAT 1092
QY 1389 AGAATTGGAGACACCTGGAGGAGAGGATATCGAGGAAACCTGTCTCCAGTGCATCTGC 1448
DB 1093 AGAATTGGAGACACCTGGAGGAGAGGATATCGAGGAAACCTGTCTCCAGTGCATCTGC 1152
QY 1449 ACAGGCAACGGCCGAGGAGAGTGAAGTGTGAGAGGACACCTCTGTGTCAGACCATCG 1508
DB 1153 ACAGGCAACGGCCGAGGAGAGTGAAGTGTGAGAGGACACCTCTGTGTCAGACCATCG 1212
QY 1509 AGCGGATCTGGCCCCCTTCACCGATGTTGCT 1538
DB 1213 AGCGGATCTGGCCCCCTTCACCGATGTTGCT 1242

RESULT 11

US-10-995-561-112
; Sequence 112, Application US/10995561
; Publication No. US20050272054A1

; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 112

```
; LENGTH: 6510
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-112
```

```
Query Match          21.4%; Score 330; DB 7; Length 6510;
Best Local Similarity 100.0%; Pred. No. 1.7e-94;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1209 CCATAGCTGAGAGAGTTTGTATCATGCTGCGGACCTTCATGTGTGCGGAAACG 1268
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 913 CCATAGCTGAGAGAGTTTGTATCATGCTGCGGACCTTCATGTGTGCGGAAACG 972

QY 1269 TGGGAGAAACCCCTACCAAGGCTGATGATGATTTGTACTTCCCTGGAGAAAGCAGC 1328
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 973 TGGGAGAAACCCCTACCAAGGCTGATGATGATTTGTACTTCCCTGGAGAAAGCAGC 1032

QY 1329 GAGCGCATCACTTGCACTTCTAGAAATAGATGCAACGATGAGACACAAAGACATCTTAT 1388
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1033 GAGCGCATCACTTGCACTTCTAGAAATAGATGCAACGATGAGACACAAAGACATCTTAT 1092

QY 1389 AGAATTGGAGACACCTGGAGCAAGAGATATCGAGAAACCTGCTCCAGTGCATCTGC 1448
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1093 AGAATTGGAGACACCTGGAGCAAGAGATATCGAGAAACCTGCTCCAGTGCATCTGC 1152

QY 1449 ACAGGCAACGCGCCGAGAGAGTGAAGTGAAGGACACCTCTGTGACAGACCATCG 1508
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1153 ACAGGCAACGCGCCGAGAGAGTGAAGTGAAGGACACCTCTGTGACAGACCATCG 1212

QY 1509 AGCGGATCTGGCCCTTCAACCGATGTTTGGT 1538
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1213 AGCGGATCTGGCCCTTCAACCGATGTTTGGT 1242
```

```
RESULT 12
US-10-995-561-117
; Sequence 117, Application US/10995561
; Publication No. US20050272054A1
```

```
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 117
; LENGTH: 7823
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-117
```

```
Query Match          21.4%; Score 330; DB 7; Length 7823;
Best Local Similarity 100.0%; Pred. No. 1.9e-94;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1209 CCATAGCTGAGAGAGTTTGTATCATGCTGCGGACCTTCATGTGTGCGGAAACG 1268
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 913 CCATAGCTGAGAGAGTTTGTATCATGCTGCGGACCTTCATGTGTGCGGAAACG 972

QY 1269 TGGGAGAAACCCCTACCAAGGCTGATGATGATTTGTACTTCCCTGGAGAAAGCAGC 1328
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 973 TGGGAGAAACCCCTACCAAGGCTGATGATGATTTGTACTTCCCTGGAGAAAGCAGC 1032

QY 1329 GAGCGCATCACTTGCACTTCTAGAAATAGATGCAACGATGAGACACAAAGACATCTTAT 1388
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1033 GAGCGCATCACTTGCACTTCTAGAAATAGATGCAACGATGAGACACAAAGACATCTTAT 1092

QY 1389 AGAATTGGAGACACCTGGAGCAAGAGATATCGAGAAACCTGCTCCAGTGCATCTGC 1448
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1093 AGAATTGGAGACACCTGGAGCAAGAGATATCGAGAAACCTGCTCCAGTGCATCTGC 1152
```

```
QY 1449 ACAGGCAACGCGCCGAGAGAGTGAAGTGAAGGACACACCTCTGTGACAGACCATCG 1508
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1153 ACAGGCAACGCGCCGAGAGAGTGAAGTGAAGGACACACCTCTGTGACAGACCATCG 1212

QY 1509 AGCGGATCTGGCCCTTCAACCGATGTTTGGT 1538
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1213 AGCGGATCTGGCCCTTCAACCGATGTTTGGT 1242
```

```
RESULT 13
US-10-995-561-111
```

```
; Sequence 111, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 111
; LENGTH: 7848
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-111
```

```
Query Match          21.4%; Score 330; DB 7; Length 7848;
Best Local Similarity 100.0%; Pred. No. 1.9e-94;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1209 CCATAGCTGAGAGAGTTTGTATCATGCTGCGGACCTTCATGTGTGCGGAAACG 1268
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 913 CCATAGCTGAGAGAGTTTGTATCATGCTGCGGACCTTCATGTGTGCGGAAACG 972

QY 1269 TGGGAGAAACCCCTACCAAGGCTGATGATGATTTGTACTTCCCTGGAGAAAGCAGC 1328
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 973 TGGGAGAAACCCCTACCAAGGCTGATGATGATTTGTACTTCCCTGGAGAAAGCAGC 1032

QY 1329 GAGCGCATCACTTGCACTTCTAGAAATAGATGCAACGATGAGACACAAAGACATCTTAT 1388
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1033 GAGCGCATCACTTGCACTTCTAGAAATAGATGCAACGATGAGACACAAAGACATCTTAT 1092

QY 1389 AGAATTGGAGACACCTGAGACAAAGAGATATCGAGAAACCTGCTCCAGTGCATCTGC 1448
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1093 AGAATTGGAGACACCTGAGACAAAGAGATATCGAGAAACCTGCTCCAGTGCATCTGC 1152

QY 1449 ACAGGCAACGCGCCGAGAGAGTGAAGTGAAGGACACACCTCTGTGACAGACCATCG 1508
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1153 ACAGGCAACGCGCCGAGAGAGTGAAGTGAAGGACACACCTCTGTGACAGACCATCG 1212

QY 1509 AGCGGATCTGGCCCTTCAACCGATGTTTGGT 1538
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1213 AGCGGATCTGGCCCTTCAACCGATGTTTGGT 1242
```

```
RESULT 14
US-10-995-561-113
```

```
; Sequence 113, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
```

```

; SEQ ID NO 113
; LENGTH: 7935
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-113

Query Match      21.4%; Score 330; DB 7; Length 7935;
Best Local Similarity 100.0%; Pred. No. 1.9e-94;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1209 CCCATAGCTGAGAAGTGTTCATCATGCTGCTGGGACCTTCCTATGTTGGTCCGAGAAACG 1268
Db 913 CCCATAGCTGAGAAGTGTTCATCATGCTGCTGGGACCTTCCTATGTTGGTCCGAGAAACG 972

QY 1269 TGGGAGAGCCCTACCAAGGCTGGATGATGTTAGTGTGCTGGGAGGAGCAGC 1328
Db 973 TGGGAGAGCCCTACCAAGGCTGGATGATGTTAGTGTGCTGGGAGGAGCAGC 1032

QY 1329 GGACGCATCATTGCACTTTAGAAATAGATCAACGATCAGGACACAAAGGACATCCTAT 1388
Db 1033 GGACGCATCATTGCACTTTAGAAATAGATCAACGATCAGGACACAAAGGACATCCTAT 1092

QY 1389 AGAATTGGAGACACCTGGAGCAAGAGGATATCGAGGAAACCTGCTCCAGTGCATCTGC 1448
Db 1093 AGAATTGGAGACACCTGGAGCAAGAGGATATCGAGGAAACCTGCTCCAGTGCATCTGC 1152

QY 1449 ACAGGCAACGGCGGAGGAGTGGAGTGTGAGAGGCACACCTCTGTGCAGACCATCG 1508
Db 1153 ACAGGCAACGGCGGAGGAGTGGAGTGTGAGAGGCACACCTCTGTGCAGACCATCG 1212

QY 1509 AGCGGATCTGGCCCTTACCCGATGTTGCT 1538
Db 1213 AGCGGATCTGGCCCTTACCCGATGTTGCT 1242

```

Search completed: February 1, 2006, 14:43:12
Job time : 285.911 secs

```

; SEQ ID NO 113
; LENGTH: 7935
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-113

Query Match      21.4%; Score 330; DB 7; Length 7935;
Best Local Similarity 100.0%; Pred. No. 1.9e-94;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1209 CCCATAGCTGAGAAGTGTTCATCATGCTGCTGGGACCTTCCTATGTTGGTCCGAGAAACG 1268
Db 913 CCCATAGCTGAGAAGTGTTCATCATGCTGCTGGGACCTTCCTATGTTGGTCCGAGAAACG 972

QY 1269 TGGGAGAGCCCTACCAAGGCTGGATGATGTTAGTGTGCTGGGAGGAGCAGC 1328
Db 973 TGGGAGAGCCCTACCAAGGCTGGATGATGTTAGTGTGCTGGGAGGAGCAGC 1032

QY 1329 GGACGCATCATTGCACTTTAGAAATAGATCAACGATCAGGACACAAAGGACATCCTAT 1388
Db 1033 GGACGCATCATTGCACTTTAGAAATAGATCAACGATCAGGACACAAAGGACATCCTAT 1092

QY 1389 AGAATTGGAGACACCTGGAGCAAGAGGATATCGAGGAAACCTGCTCCAGTGCATCTGC 1448
Db 1093 AGAATTGGAGACACCTGGAGCAAGAGGATATCGAGGAAACCTGCTCCAGTGCATCTGC 1152

QY 1449 ACAGGCAACGGCGGAGGAGTGGAGTGTGAGAGGCACACCTCTGTGCAGACCATCG 1508
Db 1153 ACAGGCAACGGCGGAGGAGTGGAGTGTGAGAGGCACACCTCTGTGCAGACCATCG 1212

QY 1509 AGCGGATCTGGCCCTTACCCGATGTTGCT 1538
Db 1213 AGCGGATCTGGCCCTTACCCGATGTTGCT 1242

```

```

RESULT 15
US-10-995-561-108
; Sequence 108, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 108
; LENGTH: 7959
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-108,

```

```

Query Match      21.4%; Score 330; DB 7; Length 7959;
Best Local Similarity 100.0%; Pred. No. 1.9e-94;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1209 CCCATAGCTGAGAAGTGTTCATCATGCTGCTGGGACCTTCCTATGTTGGTCCGAGAAACG 1268
Db 913 CCCATAGCTGAGAAGTGTTCATCATGCTGCTGGGACCTTCCTATGTTGGTCCGAGAAACG 972

QY 1269 TGGGAGAGCCCTACCAAGGCTGGATGATGTTAGTGTGCTGGGAGGAGCAGC 1328
Db 973 TGGGAGAGCCCTACCAAGGCTGGATGATGTTAGTGTGCTGGGAGGAGCAGC 1032

QY 1329 GGACGCATCATTGCACTTTAGAAATAGATCAACGATCAGGACACAAAGGACATCCTAT 1388
Db 1033 GGACGCATCATTGCACTTTAGAAATAGATCAACGATCAGGACACAAAGGACATCCTAT 1092

QY 1389 AGAATTGGAGACACCTGGAGCAAGAGGATATCGAGGAAACCTGCTCCAGTGCATCTGC 1448

```

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2006, 16:09:44 ; Search time 6332.25 Seconds
(without alignments)
11385.988 Million cell updates/sec

Title: US-09-940-235-9

Perfect score: 1541
Sequence: 1 ttgttcaacttaagaagg.....cctcacgagatgtcgttag 1541

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Maximum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	330	21.4	451	1	AL603368 DKFZp686C
2	330	21.4	560	7	CN419585 170005318
3	330	21.4	621	7	CN482442 h20d08.Y
4	330	21.4	626	7	CN419611 170005313
5	330	21.4	737	1	AU140993 AU140993
6	330	21.4	765	3	BM715855 UI-E-B70-
7	330	21.4	861	1	AU140971 AU140971
8	330	21.4	7501	4	BC078656 Homo sapi
9	330	21.4	7501	4	BC100030 Homo sapi
10	330	21.4	7777	4	CR749316 Homo sapi
11	330	21.4	7868	4	HSW80653 Homo sapi
12	330	21.4	7885	4	CR749281 Homo sapi
13	330	21.4	8411	4	HSW80692 Homo sapi
14	329	21.3	465	1	AI095589 gbd24a08.x
15	329	21.3	734	1	AU140802 AU140802
16	328.4	21.3	8121	4	CR749317 Homo sapi
17	326.8	21.2	548	6	CD613789 56022208U
18	325.2	21.1	551	6	CD613788 56022208U
19	325.2	21.1	560	1	AI743013 w95a12.x
20	319	20.7	500	2	BF956982 OVA-NN114
21	318.2	20.6	547	2	BE009640 PM4-BN017
22	318	20.6	943	5	BX391752 BX391752

23	317	20.6	705	1	AU140789
24	315.4	20.5	739	1	AL706215
25	310	20.1	603	7	CN419658
26	307	19.9	414	1	AA492032
27	305	19.8	518	1	AI093548
28	304	19.7	7434	11	DQ039102
29	298	19.3	538	1	AL603599
30	297.4	19.3	496	2	BG945197
31	296.4	19.2	303	1	AA85090
32	287	18.6	427	7	CN419650
33	287	18.6	452	7	CN419649
34	281	18.2	605	2	CN419657
35	280	18.2	408	7	BG900107
36	277.8	18.0	580	3	BQ292415
37	274	17.8	495	2	BF956977
38	271.8	17.6	494	2	BF996365
39	268.6	17.4	484	2	BI058354
40	268	17.4	725	5	BQ574857
41	264	17.1	454	8	W46530
42	261.8	17.0	474	6	CB536952
43	256.8	16.7	645	8	DN123922
44	251.8	16.3	314	7	CN419631
45	249.4	16.2	899	5	BX327266

ALIGNMENTS

RESULT 1
AL603368 451 bp mRNA linear EST 04-SEP-2003
LOCUS DKFZp686C067.5, mRNA sequence.
DEFINITION DKFZp686C067.5, mRNA sequence.
ACCESSION AL603368
VERSION AL603368.1 GI:15166874
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 451)
AUTHORS Ansoorge,W., Winkler,U., Mewes,W., Weil,B. and Wiemann,S.
TITLES EST (Ansoorge,W., Winkler,U., Mewes,W., Weil,B. and Wiemann,S.)
JOURNAL Unpublished (1999)
COMMENT Contact: MIPS

MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by EMBL (European Molecular Biology Laboratories), Heidelberg/Germany) within the CDNA sequencing consortium of the German Genome Project.
No 5' sequence available.
This clone (DKFZp686C067) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcencentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

source
1..451
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686C067"
/dev_stage="adult"
/lab_host="DH10B"
/clone_id="586 (synonym: hlc3)"
/note="Vector: pTIP1Ex2; Site_1: SfiIA; Site_2: SfiIB; cDNA-collection"

ORIGIN

Query Match 21.4%; Score 330; DB 1; Length 451;
Best Local Similarity 100.0%; Pred. No. 1.1e-82;

Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1209 CCATAGCTGAGAAGTGTGTTTGCATCATGCTGCTGGGACTTCCTATGTTGGTCGGAGAAACG 1268
|||||
Db 32 CCATAGCTGAGAAGTGTGTTTGCATCATGCTGCTGGGACTTCCTATGTTGGTCGGAGAAACG 91
|||||
QY 1269 TGGGAGAAGCCCTACCAAGGCTGGATGATGGTAGATTGTAATTGCTTGGGAGAAAGGCAGC 1328
|||||
Db 92 TGGGAGAAGCCCTACCAAGGCTGGATGATGGTAGATTGTAATTGCTTGGGAGAAAGGCAGC 151
|||||
QY 1329 GGACGCATCACTTGCACCTCTAGAAATAGATGCAACGATCAGGACACAGGACATCCTAT 1388
|||||
Db 152 GGACGCATCACTTGCACCTCTAGAAATAGATGCAACGATCAGGACACAGGACATCCTAT 211
|||||
QY 1389 AGAATTGGAGACACCTTGGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC 1448
|||||
Db 212 AGAATTGGAGACACCTTGGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC 271
|||||
QY 1449 ACAGGCAACGGCCGAGGAGTGGAGTGTGAGAGGCACACCTCTGTGCAGACCAATCG 1508
|||||
Db 272 ACAGGCAACGGCCGAGGAGTGGAGTGTGAGAGGCACACCTCTGTGCAGACCAATCG 331
|||||
QY 1509 AGCGGATCTGGCCCTTCAACCGATGTTTGGT 1538
|||||
Db 332 AGCGGATCTGGCCCTTCAACCGATGTTTGGT 361
|||||

RESULT 2
CN419585
LOCUS
DEFINITION
1700531856084 GRN_EB Homo sapiens cDNA 5', mRNA sequence.
CN419585
VERSION
CN419585.1 GI:47407179
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 (bases 1 to 560)
Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
Li,Y., Xu,C., Fang,R., Guegler,K., Rac,M.S., Mandalam,R.,
Lebkowski,J and Stanton,L.W.
Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation
Nat. Biotechnol. 22 (6), 707-716 (2004)
15146197
Contact: Brandenberger R
Regenerative Medicine
Genon Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@genon.com
Insert Length: 560 Std Error: 0.00.
FEATURES
Location/Qualifiers
1..560
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue type="embryonic stem cells, embryoid bodies
derived from H1, H7 and H9 cells"
/clone_lib="GRN EB"
/note="oligo dt primed, full-length enriched cDNA library
from embryoid body outgrowths derived from hES cell lines
H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free
conditions."
ORIGIN
Query Match 21.4%; Score 330; DB 7; Length 560;
Best Local Similarity 100.0%; Pred. No. 1.2e-82;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1209 CCATAGCTGAGAAGTGTGTTTGCATCATGCTGCTGGGACTTCCTATGTTGGTCGGAGAAACG 1268
|||||
Db 209 CCATAGCTGAGAAGTGTGTTTGCATCATGCTGCTGGGACTTCCTATGTTGGTCGGAGAAACG 268
|||||
QY 1269 TGGGAGAAGCCCTACCAAGGCTGGATGATGGTAGATTGTAATTGCTTGGGAGAAAGGCAGC 1328
|||||
Db 269 TGGGAGAAGCCCTACCAAGGCTGGATGATGGTAGATTGTAATTGCTTGGGAGAAAGGCAGC 328
|||||
QY 1329 GGACGCATCACTTGCACCTCTAGAAATAGATGCAACGATCAGGACACAGGACATCCTAT 1388
|||||
Db 329 GGACGCATCACTTGCACCTCTAGAAATAGATGCAACGATCAGGACACAGGACATCCTAT 388
|||||
QY 1389 AGAATTGGAGACACCTTGGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC 1448
|||||
Db 389 AGAATTGGAGACACCTTGGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC 448
|||||
QY 1449 ACAGGCAACGGCCGAGGAGTGGAGTGTGAGAGGCACACCTCTGTGCAGACCAATCG 1508
|||||
Db 449 ACAGGCAACGGCCGAGGAGTGGAGTGTGAGAGGCACACCTCTGTGCAGACCAATCG 508
|||||
QY 1509 AGCGGATCTGGCCCTTCAACCGATGTTTGGT 1538
|||||
Db 509 AGCGGATCTGGCCCTTCAACCGATGTTTGGT 538
|||||

RESULT 3
CN482442
LOCUS
DEFINITION
hw20d08.y1 Human primary human ocular pericytes. Unamplified (hw)
Homo sapiens cDNA clone hw20d08 5', mRNA sequence.
CN482442
VERSION
CN482442.1 GI:46563946
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 (bases 1 to 621)
Tsai,J.Y. and Wistow,G.
Expressed sequence tag analysis of cultured primary human ocular
pericytes
Unpublished (2004)
Contact: Wistow G
Section on Molecular Structure and Function
National Eye Institute
6/331, NIH, Bethesda, MD 20892-2740, USA
Tel: 301 402 3452
Fax: 301 496 0078
Email: graeme@helix.nih.gov
Plate: 20 row: d column: 08
Seq primer: M13RP1 reverse primer (ABI).
FEATURES
Location/Qualifiers
1..621
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="hw20d08"
/cell_type="pericytes"
/dev_stage="Adult"
/lab_host="EMDH10B"
/clone_lib="Human primary human ocular pericytes."
Unamplified (hw)
/note="Organ: Eye; Vector: pSPORT1; RNA was extracted from
primary human pericytes in culture. A directionally cloned
cDNA library in the pSPORT1 vector (Invitrogen) was
constructed at Bioserve Biotechnology (Laurel MD)
essentially following the protocols of the SuperScript
Plasmid System full details of which are contained in the
manufacturer's instruction manual
(http://www.lifetech.com/). First strand synthesis was
carried out using a Not I primer-adaptor
[5'-pGACTAGTCTTAGATCGGCGCGGCC(T)15-3']. cDNA was


```
Query Match      21.4%; Score 330; DB 1; Length 737;
Best Local Similarity 100.0%; Pred. No. 1.3e-82;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1209 CCCATAGCTGAGAAGTGTGTTGATCATGCTGCTGGGACTTCCTATGTTGGTCGGAGAAACG 1268
Db 362 CCCATAGCTGAGAAGTGTGTTGATCATGCTGCTGGGACTTCCTATGTTGGTCGGAGAAACG 421

QY 1269 TGGGAGAACCCCTACCAAGGCTGGATGATGTTGATGTTGCTTGGGAGAAAGGCAGC 1328
Db 422 TGGGAGAACCCCTACCAAGGCTGGATGATGTTGATGTTGCTTGGGAGAAAGGCAGC 481

QY 1329 GGAAGCATCACTTGCCTTCTAGAAATAGATCAACGATCAGGACACAGGACATCCTAT 1388
Db 482 GGAAGCATCACTTGCCTTCTAGAAATAGATCAACGATCAGGACACAGGACATCCTAT 541

QY 1389 AGAATTGGAGACACCTGGAGCAAGAGGATTAATCGAGGAAACCTGCTCCAGTGCATCTGC 1448
Db 542 AGAATTGGAGACACCTGGAGCAAGAGGATTAATCGAGGAAACCTGCTCCAGTGCATCTGC 601

QY 1449 ACAGGCAACGGCGGAGGAGTGGAAAGTGTGAGAGGACACACCTCTGTGACAGACCATCG 1508
Db 602 ACAGGCAACGGCGGAGGAGTGGAAAGTGTGAGAGGACACACCTCTGTGACAGACCATCG 661

QY 1509 AGCGGATCTGGCCCTTCCACCGATGTTTGGT 1538
Db 662 AGCGGATCTGGCCCTTCCACCGATGTTTGGT 691

RESULT 6
BM715855
LOCUS      BM715855      765 bp      mRNA      linear      EST 28-FEB-2002
DEFINITION UI-E-EJ0-abj-h-11-0-UI.r2 UI-E-EJ0 Homo sapiens cDNA clone
ACCESSION  BM715855
VERSION     BM715855.1 GI:19029113
KEYWORDS   Homo sapiens (human)
SOURCE     Homo sapiens
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE  1 (bases 1 to 765)
AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE     Normalization and subtraction: two approaches to facilitate gene
          discovery
JOURNAL   Genome Res. 6 (9), 791-806 (1996)
PUBMED   8889548
COMMENT   Contact: Soares, MB
          Coordinated Laboratory for Computational Genomics
          University of Iowa
          375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA
          Tel: 319 335 8250
          Fax: 319 335 9565
          Email: bento-soares@uiowa.edu
          Tissue Procurement: Dr. Gregg Hageman
          cDNA library preparation: Dr. M. Bento Soares, University of Iowa
          cDNA library arrayed by: Dr. M. Bento Soares, University of Iowa
          DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
          Clone Distribution: Researchers may obtain clones from Research
          Genetics (www.resgen.com).
          Seq primer: M13 Reverse
          Location/Qualifiers
            1..765
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="UI-E-EJ0-abj-h-11-0-UI"
              /tissue_type="fetal eyes, lens, eye anterior segment,
              optic nerve, retina, Retina Foveal and Macular, RPE and
              Choroid"
              /dev_stage="fetal and adult"
              /lab_host="DH10B (Life Technologies) (T1 phage resistant)"

FEATURES
source
```

```
/clone lib="UI-E-EJ0"
/notes="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-E-EJ0 is a subcloned cDNA library constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pT73-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tags for this library are: fetal eyes,
AGATCAACA; lens, CGATTAGCGA; eye anterior segment,
AATCCGCGAT; optic nerve, CCATTAGTGT; retina, CGCG; Retina
Foveal and Macular, GTCC; RPE and Choroid, ACCTA. This
library was created for the program, Gene Discovery in the
Visual System, supported by National Eye Institute (NEI)."
```

ORIGIN

Query Match 21.4%; Score 330; DB 3; Length 765;
Best Local Similarity 100.0%; Pred. No. 1.3e-82;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1209 CCCATAGCTGAGAAGTGTGTTGATCATGCTGCTGGGACTTCCTATGTTGGTCGGAGAAACG 1268
Db 182 CCCATAGCTGAGAAGTGTGTTGATCATGCTGCTGGGACTTCCTATGTTGGTCGGAGAAACG 241

QY 1269 TGGGAGAACCCCTACCAAGGCTGGATGATGTTGATGTTGCTTGGGAGAAAGGCAGC 1328
Db 242 TGGGAGAACCCCTACCAAGGCTGGATGATGTTGATGTTGCTTGGGAGAAAGGCAGC 301

QY 1329 GGAAGCATCACTTGCCTTCTAGAAATAGATCAACGATCAGGACACAGGACATCCTAT 1388
Db 302 GGAAGCATCACTTGCCTTCTAGAAATAGATCAACGATCAGGACACAGGACATCCTAT 361

QY 1389 AGAATTGGAGACACCTGGAGCAAGAGGATTAATCGAGGAAACCTGCTCCAGTGCATCTGC 1448
Db 362 AGAATTGGAGACACCTGGAGCAAGAGGATTAATCGAGGAAACCTGCTCCAGTGCATCTGC 421

QY 1449 ACAGGCAACGGCGGAGGAGTGGAAAGTGTGAGAGGACACACCTCTGTGACAGACCATCG 1508
Db 422 ACAGGCAACGGCGGAGGAGTGGAAAGTGTGAGAGGACACACCTCTGTGACAGACCATCG 481

QY 1509 AGCGGATCTGGCCCTTCCACCGATGTTTGGT 1538
Db 482 AGCGGATCTGGCCCTTCCACCGATGTTTGGT 511

RESULT 7
AU140971
LOCUS AU140971 PLACE4 Homo sapiens cDNA clone PLACE4000583 5', mRNA
DEFINITION sequence.
ACCESSION AU140971
VERSION AU140971.1 GI:11002492
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 861)
AUTHORS Ota,T., Suzuki,Y., Saito,K., Ishii,S., Yamamoto,J., Sugiyama,T.,
 Nishikawa,T., Nakamura,Y., Sugano,S., Masuho,Y. and Isogai,T.
 HRI human cDNA project (Ota,T., Suzuki,Y., Saito,K., Ishii,S.,
 Yamamoto,J., Sugiyama,T., Nishikawa,T., Nakamura,Y., Sugano,S.,
 Masuho,Y., Isogai,T.)
 Unpublished (2000)
 Contact: Takao Isogai
 Genomics Laboratory
 Helix Research Institute
 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan

Tel: 81-438-52-3975
 Fax: 81-438-52-3986
 Email: genomics@hri.co.jp
 HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
 Research Institute; cDNA library construction: Department of
 Virology, Institute of Medical Science, University of Tokyo, and
 Helix Research Institute.

FEATURES

source
 1. 861
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="PLACE4000583"
 /tissue_type="placenta"
 /clone_lib="PLACE4"
 /note="Vector: pME18SFL3"

ORIGIN

Query Match 21.4%; Score 330; DB 1; Length 861;
 Best Local Similarity 100.0%; Pred. No. 1.4e-82;
 Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1209 CCCATGCTGAGAGAGTTTGTATCATCTGCTGGAGCTTCTATGTGTGGAGAAACG 1268
 |||||
 DB 158 CCCATGCTGAGAGAGTTTGTATCATCTGCTGGAGCTTCTATGTGTGGAGAAACG 217
 |||||
 QY 1269 TGGGAGAGAGCCCTACAGAGCTGGATGATGTGTATGTTCTGCTGGAGAGAGCAGC 1328
 |||||
 DB 218 TGGGAGAGAGCCCTACAGAGCTGGATGATGTGTATGTTCTGCTGGAGAGAGCAGC 277
 |||||
 QY 1329 GGAGCGATCATCTGCACTTCTAGAAATAGATGCAACGATCAGACACAGAGACATCTAT 1388
 |||||
 DB 278 GGAGCGATCATCTGCACTTCTAGAAATAGATGCAACGATCAGACACAGAGACATCTAT 337
 |||||
 QY 1389 AGAATTGGAGACACTCTGAGCAAGAGATATCGAGAAACCTGCTCCAGTGCATCTGC 1448
 |||||
 DB 338 AGAATTGGAGACACTCTGAGCAAGAGATATCGAGAAACCTGCTCCAGTGCATCTGC 397
 |||||
 QY 1449 AAGAGCAAGCCCGGAGAGAGTGGAGTGTGAGAGGCAACCTCTGTGCGAGACACATCG 1508
 |||||
 DB 398 AAGAGCAAGCCCGGAGAGAGTGGAGTGTGAGAGGCAACCTCTGTGCGAGACACATCG 457
 |||||
 QY 1509 AGCGGATCTGGCCCTTCAACCGATGTTCCG 1538
 |||||
 DB 458 AGCGGATCTGGCCCTTCAACCGATGTTCCG 487
 |||||

RESULT 8
 BC078656 7501 bp mRNA linear HTC 21-JUL-2005
 LOCUS Homo sapiens cDNA clone IMAGE:30347017, containing frame-shift
 errors.
 DEFINITION BC078656
 VERSION BC078656.1 GI:50925326
 KEYWORDS HTC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homiidae; Homo.
 1 (bases 1 to 7501)

REFERENCE
 AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Haib, F.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
 Carrinzi, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullany, S.J., Bosnak, S.A., McGwan, P.J.,
 McKernan, K.J., Malek, J.A., Gamarall, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hult, S.W.,
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,

Fahy, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butcherfield, Y.S., Krzywnicki, M.I., Skalska, U., Smalins, D.E.,
 Schnerch, A., Schein, J.E., Jones, S.J., and Warr, M.A.
 Mammalian Gene Collection Program Team
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

ORIGIN

Query Match 21.4%; Score 330; DB 4; Length 7501;
 Best Local Similarity 100.0%; Pred. No. 2.5e-82;
 Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1209 CCCATGCTGAGAGAGTTTGTATCATCTGCTGGAGCTTCTATGTGTGGAGAAACG 1268
 |||||
 DB 806 CCCATGCTGAGAGAGTTTGTATCATCTGCTGGAGCTTCTATGTGTGGAGAAACG 865
 |||||
 QY 1269 TGGGAGAGAGCCCTACAGAGCTGGATGATGTGTATGTTCTGCTGGAGAGAGCAGC 1328
 |||||
 DB 866 TGGGAGAGAGCCCTACAGAGCTGGATGATGTGTATGTTCTGCTGGAGAGAGCAGC 925
 |||||
 QY 1329 GGAGCGATCATCTGCACTTCTAGAAATAGATGCAACGATCAGACACAGAGACATCTAT 1388
 |||||
 DB 926 GGAGCGATCATCTGCACTTCTAGAAATAGATGCAACGATCAGACACAGAGACATCTAT 985
 |||||
 QY 1389 AGAATTGGAGACACTCTGAGCAAGAGATATCGAGAAACCTGCTCCAGTGCATCTGC 1448
 |||||
 DB 986 AGAATTGGAGACACTCTGAGCAAGAGATATCGAGAAACCTGCTCCAGTGCATCTGC 1045
 |||||
 QY 1449 ACAAGCAAGCCCGGAGAGAGTGGAGTGTGAGAGGCAACCTCTGTGCGAGACACATCG 1508
 |||||
 DB 1046 ACAAGCAAGCCCGGAGAGAGTGGAGTGTGAGAGGCAACCTCTGTGCGAGACACATCG 1105
 |||||

FEATURES

source
 1. 7501
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30347017"
 /tissue_type="placenta, normal"
 /clone_lib="NIH_MGC_147"
 /lab_host="DH10B"
 /note="Vector: pBluescriptR"

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LINL at: <http://image.llnl.gov>
 Series: IRAC Plate: 168 Row: 0 Column: 20
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 47132556
 This clone has the following problem: frame shifted.

gene
CDS

```

/clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host
DH10B; sites SfiIA + SfiIB"
/dev_stage="adult"
/notes="fibronectin 1, differentially spliced"
1..7777
/gene="DKFZp686K0164"
/266..7069
/gene="DKFZp686K0164"
/codon_start=1
/product="hypothetical protein"
/db_xref="GI:51476362"
/translation="MLRGPGLLLAVLCGLAVBSTASKSKRRAQMVQPSVA
VSQSKPCYDNGKHQVINOQWERTYLGMLVCTCYGSGRFNCSKREAEETFDKXT
GNTYVGVGYERPKDSIMWDCICIGAGRIISCTIANGRCHGSGYKIGIAPRPHET
GGYMLCVCCLGNKGEWTKCKPIAEKCFDHAAGSYVVGTEWKEPQGMVNDCTGGE
GSGRLTCSRNRCNDODTPRSYRIGDTWSKDRGNLLOCTICGNRGKRCRHYSV
QTTSSGSGPFTVRAAVYOPPHPPPPGGHCVTDSGVVSVGMQMLKTOGNKQMLCT
CLGNVSCQETAVTQTYGNSNGEPCVLPFTYNGRTYSCCTEGRDGHLWCSTSNY
BODQKYSFCTDHTVLTQYTGNSNGALCHPPLVNNHNTYDCTSEGRDNMKKQCTVO
NYDADQKFGCPMAAHESICTNKGWYRIGDQMDKQMDHGMHMRCTCGNGRGEWTO
IAYSQLDQCIYDITVYNDVPHKREHGHMLNCTCGOGRGRKCDPDDOCSDET
GTFYQIGSKRYKHCYRQCYCYGGRIGEMHQPIQTPSSSGPVPVITTEPSPN
SHIPQMANAPQSHISKYILKMPKNSVGWKEKATLGHLSNTYIKLKGVTVEGLI
SIQYQGHQVTRFDFTTSTSPVTSNTVGTETTPSPPLVASSESTETASFSVSM
VSADTVSGFRVYELSEBDEPOYLDLPSTASVNIPLLPKRTYVWVQISGESE
OSLSTSQETAPADAPPDTPVQDQDTSIVWMSRQAPLGTGRYIVSPSSESTEL
NLEPANSVTLMDQGVQYNTIYAVENOSTPVOEOTETGTPRSPVPSRLQ
PFEVTVDKYTIWMTTPRESAVTGRVDVITVNLPGEGQGLPISKRTPAEYTGLSPEVT
YFKVFAVSHGRESKPLTAQOTTKLDAFNLQVNSTDSTVLAWRMPLAQITGYALT
VGLTRGQRPQYNVGSVSKYPLRMNIGPSESVYSVAIKGNQESKACVFTTLQPG
SSIPYNEVEETTVITWTTPAIRPKLGRVPSOGGEARVTSOGSIVSGTLQPG
VEYVYTIQVLRGQERDIAVNVKVTPLSPPTNHLAMPDQGLTVMSERSTPDT
GRITTPNNGOQNSLEAVNADOSSCTPDLMSBELNVSYYTKDKESPIEDT
IIPVQULTDLSFVIDTSSIGKRTPLNSSTIIGRTVVAAGSEPIPEPVDSSV
GYVTYGLRPGIDYDISVTLLNGGESAPITTLQOAVAPPDPLRNTNIGPDMRTAT
APPSIDLTNPLFVRSVKNREDAVALISPSDNVAVLNLPGTYVVSVSZYOH
ESTPLRGOKTGLDSTGIDFSDITFANSFTVMIAPRATITGRIRHHPHESGRRE
DRVPSRNSITLNLTPGREVYSIVANGRESPLIGOGSRKPLDTPVADPVTAVATP
TSLISMDAPATVAYRITTEGETGNSVQETTPGASATISLAKGVDTITTV
AVTGRGDSPPASKPISINRTIEDPSQNVQVNDONSISVWLPSSSVYGRVATTT
PRNGPPTKTKTAGPDQEMTIEGLQPYVEVYVVAQPSGQPLVOTAVTITAP
DLKFTQVPTLSLAQMTPPNQLTGRVATPKETGEMKEINLAPDSVVVSGLM
VATKEVSVYALKDILTSRPAQGVTLLENVSPPRARVTDATETTTISWTKETI
TGFOVDAPANGOTPIORTIKPDSVYITGQCPGDYKLYATLNDARSSPVIDA
STAIADPBNRFLATTPNSLVSWPPRARIITGILIKESPPREVPKRPQVTE
ATTGLBETETIIVIALKNQKSEPLIGRKTQGEALQTTISWAPQDPTSEYIIS
CHPVGTDEPDLQFRPVGTSTATLGLTGTGATNTIIVELKDOQRKRVREEVTVNS
VNBGLNQPETDSCFPYTVSHYAVGDEMREMSGFKLCCGCLGFSGSGRCDSSRMC
HNGVNYKIGEKMDRGNGENOMMSCTCLGNKGEPKCDPHEATCYDNDGTYVYGEQMO
KEYLGATCSCTGCGGRCNCRPRGSEPTPEGTGGSYNOYSORHYORNTVANC
PIECFMPJDVQADREDSRE"

```

ORIGIN

Query Match 21.4%; Score 330; DB 4; Length 7777;
 Best Local Similarity 100.0%; Pred. No. 2.5e-82;
 Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1309 CCATAGCTGAGAAGTGTTCATCATGCTGCGGACTTCATGTCGAGAAACG 1268
 Db CCCATAGCTGAGAAGTGTTCATCATGCTGCGGACTTCATGTCGAGAAACG 865
 1269 TGGGGAAGCCCTACCAAGGCTGATGATGTTGATCTTCTGCGAAGAGCAGC 1328
 Db TGGGGAAGCCCTACCAAGGCTGATGATGTTGATCTTCTGCGAAGAGCAGC 925
 1329 GGAAGCACTTGAATCTTGAATGAAATAGTGAACATCAGACACATCTTAT 1388
 Db GGAAGCACTTGAATCTTGAATGAAATAGTGAACATCAGACACATCTTAT 985
 1389 AGAATTGAGACACCTGAGACAAAGATTAATCGAAGAACTGCTCCAGTGCATCTGC 1448
 Db AGAATTGAGACACCTGAGACAAAGATTAATCGAAGAACTGCTCCAGTGCATCTGC 1045

QY 1449 ACAGGCAACGCGCAGAGAGAGTGAAGTGTGAGAGACACCTCTGTGACAGACATCG 1508
 Db 1046 ACAGGCAACGCGCAGAGAGAGTGAAGTGTGAGAGACACCTCTGTGACAGACATCG 1105
 QY 1509 AGCGGATCTGGCCCTTCACCGATGTTGCT 1538
 Db 1106 AGCGGATCTGGCCCTTCACCGATGTTGCT 1135

RESULT 11

HSMB06653
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

REFERENCE
 AUTHORS
 CONSRM
 TITLE
 JOURNAL

COMMENT

This clone (DKFZp686M04163) is available at the RZPD Deutsches
 Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
 Please contact RZPD for ordering:
 http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneid=DKFZp686M04163
 Further information about the clone and the sequencing project is
 available at http://mips.gsf.de/projects/cdna/.

FEATURES

source

/organism="Homo sapiens"
 /mol_type="RNA"
 /db_xref="RZPD:DKFZp686M04163Q"
 /db_xref="taxon:9606"
 /clone="DKFZp686M04163"
 /issue_type="uterus endothel, primary cell culture"
 /clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host
 DH10B; sites SfiIA + SfiIB"
 /dev_stage="adult"
 /note="fibronectin precursor"
 1..7868
 /gene="DKFZp686M04163"
 267..7157
 /gene="DKFZp686M04163"
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="CA54714.1"
 /db_xref="GI:34364617"
 /db_xref="GCA:Q6N0A6"
 /db_xref="InterPro:IPR000083"
 /db_xref="InterPro:IPR000562"
 /db_xref="InterPro:IPR02086"
 /db_xref="InterPro:IPR003961"
 /db_xref="InterPro:IPR006209"
 /db_xref="UniProt/TREMBL:Q6N0A6"
 /translation="MLRGPGLLLAVLCGLAVBSTASKSKRRAQMVQPSVA
 VSQSKPCYDNGKHQVINOQWERTYLGMLVCTCYGSGRFNCSKREAEETFDKXT
 GNTYVGVGYERPKDSIMWDCICIGAGRIISCTIANGRCHGSGYKIGIAPRPHET
 GGYMLCVCCLGNKGEWTKCKPIAEKCFDHAAGSYVVGTEWKEPQGMVNDCTGGE
 GSGRLTCSRNRCNDODTPRSYRIGDTWSKDRGNLLOCTICGNRGKRCRHYSV
 QTTSSGSGPFTVRAAVYOPPHPPPPGGHCVTDSGVVSVGMQMLKTOGNKQMLCT

gene

CDS

Qy	1209	CCCATAGCTGAGAAGTGTTTTGTATCATGCTGCTGGGACATTCCTATGTGTGTCGAGAAACG	1268
Db	808	CCCATAGCTGAGAAGTGTTTTGTATCATGCTGCTGGGACATTCCTATGTGTGTCGAGAAACG	867
Qy	1269	TGGGAGAAGCCCTACCAAGCGTGGATGATGGTAGATTGTACTTTCCTCGGGAGAAGCGACG	1328
Db	868	TGGGAGAAGCCCTACCAAGCGTGGATGATGGTAGATTGTACTTTCCTCGGGAGAAGCGACG	927
Qy	1329	GGACGCATCACTTTCGCATCTTCAGAAAATAGATAGCAACGATCAGGACACAAAGACATCCTAT	1388
Db	928	GGACGCATCACTTTCGCATCTTCAGAAAATAGATAGCAACGATCAGGACACAAAGACATCCTAT	987
Qy	1389	AGAATTGGAGACACCTGGAGCAAGACGATAAATCGAGAAAACCTGCTCCAGTGCATCTGC	1448
Db	988	AGAATTGGAGACACCTGGAGCAAGACGATAAATCGAGAAAACCTGCTCCAGTGCATCTGC	1047
Qy	1449	ACAGGCAACGGCCGAGGAGAGTGGAAAGTGTGAGAGGCACACCTCTGTGCAGACCAATCG	1508
Db	1048	ACAGGCAACGGCCGAGGAGAGTGGAAAGTGTGAGAGGCACACCTCTGTGCAGACCAATCG	1107
Qy	1509	AGCGGATCTGGCCCCCTTTCACCGATGTTTCGT	1538
Db	1108	AGCGGATCTGGCCCCCTTTCACCGATGTTTCGT	1137

RESULT 14
AI095589/c
LOCUS
DEFINITION
AI095589 465 bp mRNA linear EST 05-OCT-1998
gb24408.x1 Soares_pregnant_uterus_NbHPU Homo sapiens cDNA clone
IMAGE:1697174 3' similar to gb:X02761_cds1 FIBRONECTIN PRECURSOR
(HUMAN);, mRNA sequence.
AI095589
AI095589.1 GI:3434565
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
1 (bases 1 to 465)
NCI-CCGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1871 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 446.
Location/Qualifiers
1. .465
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1697174"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Soares_pregnant_uterus_NbHPU"
/note="Organ: uterus; Vector: p7T73-Pac; Site 1: Not I;
Site 2: Eco RI; 1st strand cDNA was primed with a Not I -
oligo(dT) primer [5',
AATCGGAGATTCGGCGCCGCTTTTCTTTTCTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified p7T73 vector. Library
went through one round of normalization. Library
constructed by M. Fatima Bonaldo."

Query Match 21.3%; Score 329; DB 1; Length 465;

	Best Local Similarity	99.7%;	Pred. No. 2.2e-82;	
	Matches	329;	Conservative	0; Mismatches 1; Indels 0; Gaps 0;
Qy	1209	CCCATAGCTCAGAAAGTGT	TTTTGATCATGCTGCTGGGACATTCCTATGTGTGGTCGAGAAACG	1268
Db	434	CCCATAGCTCAGAAAGTGT	TTTTGATCATGCTGCTGGGACATTCCTATGTGTGGTCGAGAAACG	375
Qy	1269	TGGGAGAAGCCCTACCAAGGCTGATGATCGTATGATTTGATCTTGCCCTGGGAGAAGGCAGC	1328	
Db	374	TGGGAGAAGCCCTACCAAGGCTGATGATCGTATGATTTGATCTTGCCCTGGGAGAAGGCAGC	315	
Qy	1329	GGACGCATCACTTTGCACCTTCTAGAAAATAGATGCACACGATCAGGACACAAAGGACATCCTAT	1388	
Db	314	GGACGCATCACTTTGCACCTTCTAGAAAATAGATGCACACGATCAGGACACAAAGGACATCCTAT	255	
Qy	1389	AGAAATTGGAGACACCTGGAGCAAGAAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC	1448	
Db	254	AGAAATTGGAGACACCTGGAGCAAGAAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC	195	
Qy	1449	ACAGGCCAACGGCCGAGGAGAGTGGAAGTGTGAGAGGCACACCTTGTGTGCAGACCAACATCG	1508	
Db	194	ACAGGCCAACGGCCGAGGAGAGTGGAAGTGTGAGAGGCACACCTTGTGTGCAGACCAACATCG	135	
Qy	1509	AGCGGATCTGGGCCCTTTCACCGATGTTTCGT	1538	
Db	134	AGCGGATCTGGGCCCTTTCACCGATGTTTCGT	105	

RESULT	15
AU140802	
LOCUS	734 bp mRNA linear EST 05-AUG-2002
DEFINITION	AU140802 PLACE4 Homo sapiens cDNA clone IMAGE4000266 5', mRNA sequence.
ACCESSION	AU140802
VERSION	AU140802.1 GI:11002323
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominoidea; Hominidae; Homo. 1 (bases 1 to 734) Ota,T., Suzuki,Y., Saito,K., Ishii,S., Yamamoto,J., Sugiyama,T., Nishikawa,T., Nakamura,Y., Sugano,S., Masuo,Y. and Isogai,T. HRI human cDNA project (Ota,T., Suzuki,Y., Saito,K., Ishii,S., Yamamoto,J., Sugiyama,T., Nishikawa,T., Nakamura,Y., Sugano,S., Masuo,Y., Isogai,T.) Unpublished (2000) Contact: Takao Isogai
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
COMMENT	

```
FEATURES
source
Location/Qualifiers
1. .734
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACE4000266"
/tissue_type="placenta"
/clone_lib="PLACE4"
/note="Vector: pMR18SPT3"
```

Query Match	21.3%	Score 329;	DB 1;	Length 734;
Best Local Similarity	99.7%	Pred. No. 2.5e-82;		
Matches 329;	Conservative	0;	Mismatches 1;	Indels 0;
Gaps 0;				

```

QY 1209 CCTATAGCTGAGAAAGTTTGTATCATGCTGCTGGACTTCTATGTGTGGAGAAACG 1268
    |||||
Db 309 CCTATAGCTGAGAAAGTTTGTATCATGCTGCTGGACTTCTATGTGTGGAGAAACG 368
    |||||
QY 1269 TGGGAGAAAGCCCTTACCAAGGCTGGATGATGTGTATGATTGTACTTGCTGGAGAAAGCAGC 1328
    |||||
Db 369 TGGGAGAAAGCCCTTACCAAGGCTGGATGATGTGTATGATTGTACTTGCTGGAGAAAGCAGC 428
    |||||
QY 1329 GGACGCATCATTTGCACTTCTAGAAATPAGATGCAAGATCAGGACACAGAACATCTCTAT 1388
    |||||
Db 429 GGACGCATCATTTGCACTTCTAGAAATPAGATGCAAGATCAGGACACAGAACATCTCTAT 488
    |||||
QY 1389 AGAATTGGAGACACTTGGAGCAAGAGATTAATCGAGAAACCTGCTCCAGTGCACTTGC 1448
    |||||
Db 489 AGAATTGGAGACACTTGGAGCAAGAGATTAATCGAGAAACCTGCTCCAGTGCACTTGC 548
    |||||
QY 1449 ACAAGCAATGGCCGAGAGAGATGGAAGTGTAGAGGCAACCTCTGTGCAGACCAATCG 1508
    |||||
Db 549 ACAAGCAATGGCCGAGAGAGATGGAAGTGTAGAGGCAACCTCTGTGCAGACCAATCG 608
    |||||
QY 1509 AGCGGATCTGGGCCCTTCACCGATGTTGGT 1538
    |||||
Db 609 AGCGGATCTGGGCCCTTCACCGATGTTGGT 638
    |||||

```

Search completed: February 1, 2006, 12:22:47
 Job time : 6336.25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 28, 2006, 01:47:34 ; Search time 107.195 Seconds
(without alignments)
12632.698 Million cell updates/sec

Title: US-09-940-235-9

Perfect score: 2716

Sequence: 1 ttgttctaacttaagaag.....ccttcacgagtcgtcgttag 1541

Scoring table:

BLOSUM62	
Xgapop 10.0	Xgapext 0.5
Ygapop 10.0	Ygapext 0.5
Fgapop 6.0	Fgapext 7.0
Delop 6.0	Delext 7.0

Searched: 2443163 segs, 439378781 residues

Total number of hits satisfying chosen parameters: 4886326

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

MODEL=frame+ n2p.model -DEV=xlp
-O=/cgn2_1/USFTO_spool_p/US09940235/runat_27012006_144217_27548/app.query.fasta_1.7708
-DB=A.GeneSeq -OFMT=faстан -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09940235 @CGN_1_1_846 @runat_27012006_144217_27548 -NCPU=6 -ICPU=3
-NO_MMAP -LANG=JQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

A_GeneSeq_21:*
1: geneSeqp1980s:*
2: geneSeqp1990s:*
3: geneSeqp2000s:*
4: geneSeqp2001s:*
5: geneSeqp2002s:*
6: geneSeqp2003as:*
7: geneSeqp2003bs:*
8: geneSeqp2004s:*
9: geneSeqp2005s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1984	73.0	413	8	ADL92189
2	1984	73.0	414	2	AAR10194
3	1984	73.0	414	2	AAR63120
4	1984	73.0	414	2	AAW24794
5	1984	73.0	414	2	AAW94664
6	1984	73.0	414	3	AAV01556
7	1984	73.0	795	2	AAV4797
8	1984	73.0	795	2	AAV4797
9	1981	72.9	531	2	AAR1829

10	1980	72.9	414	2	AAW94665	AAW94665 Streptococ
11	1978	72.8	483	2	AAR12522	AAR12522 Factor Xa
12	1966.5	72.4	813	2	AAW21728	AAW21728 Wild type
13	1966.5	72.4	1194	2	AAW21726	AAW21726 Streptoki
14	1963	72.3	415	2	AAW93425	AAW93425 Streptoco
15	1958	72.1	414	2	AAW86143	AAW86143 Streptoki
16	1958	72.1	414	3	AAW01295	AAW01295 Wild type
17	1958	72.1	440	2	AAR12889	AAR12889 Streptoki
18	1958	72.1	483	3	AAR12885	AAR12885 Factor Xa
19	1958	72.1	499	2	AAR12891	AAR12891 Streptoki
20	1958	72.1	859	2	AAR12893	AAR12893 OmpA
21	1954	71.9	414	2	AAR20202	AAR20202 S.equisim
22	1951.5	71.9	415	3	AAV50870	AAV50870 Streptoco
23	1951.5	71.9	415	3	AAV95953	AAV95953 Streptoco
24	1947	71.7	414	2	AAW86144	AAW86144 De-immuni
25	1947	71.7	414	3	AAW01296	AAW01296 Altered s
26	1942	71.5	440	2	AAR06377	AAR06377 Streptoki
27	1937.5	71.3	1194	2	AAW21725	AAW21725 Modified
28	1936.5	71.3	1194	2	AAW21724	AAW21724 Modified
29	1928	71.0	413	2	AAV25020	AAV25020 Streptoki
30	1928	71.0	413	5	ABW80012	ABW80012 Streptoki
31	1928	71.0	413	6	ABG74199	ABG74199 Represent
32	1924	70.8	372	2	AAR10200	AAR10200 Streptoki
33	1916	70.5	372	2	AAR10197	AAR10197 Streptoki
34	1907.5	70.2	371	2	AAR10195	AAR10195 Streptoki
35	1907	70.2	374	2	AAR10198	AAR10198 Streptoki
36	1904	70.1	414	4	AAW74940	AAW74940 Mutant st
37	1902	70.0	384	3	AAW84007	AAW84007 Amino aci
38	1900	70.0	800	2	AAW21723	AAW21723 Plasmidog
39	1900	70.0	1181	2	AAW21727	AAW21727 Streptoki
40	1898	69.9	372	3	AAW84006	AAW84006 Amino aci
41	1890.5	69.6	747	2	AAR12894	AAR12894 Met-core
42	1887	69.5	401	3	AAW84004	AAW84004 Amino aci
43	1887	69.5	413	3	AAW84005	AAW84005 Amino aci
44	1875	69.0	369	2	AAR12892	AAR12892 Truncated
45	1855	68.3	440	1	AAW50620	AAW50620 Sequence

ALIGNMENTS

RESULT 1
ADL92189
ID ADL92189 standard; protein; 413 AA.

AC ADL92189;
XX 20-MAY-2004 (first entry)
XX Streptokinase protein sequence.
XX
XX harvesting; recombinant; host cell; N-terminal leader peptide;
XX pre-peptide; lantibiotic; post-translational modification;
XX pharmaceuticals; vaccine; immunogenic.
XX
XX OS
XX Unidentified.
XX WO2003099862-A1.
XX
XX 04-DEC-2003.
XX PD
XX 26-MAY-2003; 2003WO-NL000389.
XX PF
XX 24-MAY-2002; 2002EP-00077060.
XX PR 07-FEB-2003; 2003US-00360101.
XX
XX (NANO-) APPLIED NANOSYSTEMS BV.
XX Moll GN, Leenhouts CJ, Kuipers OP, Driessen AJM;
XX WPI; 2004-042770/04.
XX
XX Harvesting a desired polypeptide produced by a recombinant host cell, for
XX producing pharmaceuticals, comprises selecting a recombinant nucleic acid

comprising nucleic acid fragments encoding a leader peptide and the polypeptide.

Claim 4; Page 82-83; 109pp; English.

The invention relates to a novel method for harvesting a (poly)peptide produced by a recombinant host cell. The novel method involves selecting a cell comprising a first nucleic acid encoding a leader peptide and a second nucleic acid fragment encoding the desired (poly)peptide. The first and second fragments are within the same open reading frame of the first nucleic acid and the leader peptide is functionally equivalent to an N-terminal leader peptide found with the pre-peptide of a lantibiotic. The host cells and nucleic acids are useful for producing, harvesting and post-translational modification of polypeptides. The polypeptides may be used in the production of pharmaceuticals, e.g. as antigen for vaccine or immunogenic composition. This sequence represents a polypeptide relating to the novel method of the invention.

Sequence 413 AA;

Alignment Scores:

Pred. No.: 1.13e-178 Length: 413
Score: 1984.00 Matches: 383
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 73.05% Indels: 0
DB: 8 Gaps: 0

US-09-940-235-9 (1-1541) x ADL92189 (1-413)

QY	33	ATAGCTGCTCTGGAATGGCTACTAGATCGCTCTCTGTAATAACAGCCAAATGGTGT	92
DB	1	IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal	20
QY	93	AGCGTGTGCTGCTGTTGGAGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAATC	152
DB	21	SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPheGluIle	40
QY	153	GATCTAAATCATCGACCTGCTCATGGAGGAAAGACAGACAGGCTTAAGTCCAAATCA	212
DB	41	AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer	60
QY	213	AAACCATTTGCTACTGATGTCGCGATGTCATACATAAACTTGAGAAAGCTGACTACTA	272
DB	61	LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeu	80
QY	273	AAGGCTATTCAAGAACAAFTGATCGCTAACGTCACAGTAACGACGACTACTTTGAGGTC	332
DB	81	LysAlalleGlnGluLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal	100
QY	333	ATTGATTTTGCAGGATGCAACCAATTACTGATCGAAACGGCAGGTCTACTTTGCTGAC	392
DB	101	IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp	120
QY	393	AAAGATGTCGTCGTAACCTTCGCGACCCCACTGTCCTCAAGAAATTTTGTAAAGCGGACAT	452
DB	121	LysaspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis	140
QY	453	GTGCGCTTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTGTGATGTG	512
DB	141	ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal	160
QY	513	GAATATCTGTACGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGCAGGCTCTCAA	572
DB	161	GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys	180
QY	573	GATATCAAGCTATTCAAAACACTAGCTATCGGTGACCATCACATCTCAAGAAATCTA	632
DB	181	AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu	200
QY	633	GCTCAAGCAGCAAGCATTTTAAACAAAAACCCAGGCTATACGATTTTATGAACCTGAC	692
DB	201	AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp	220

QY	693	TCCTCAATCGTCACTCATGCAATGACATTTTCGTACGATTTTACCAATGGATCAAGAG	752
DB	221	SerSerIleValThrHisaspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu	240
QY	753	TTTACTTACCGTGTAAAAATCGGAAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG	812
DB	241	PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu	260
QY	813	AATGAAGAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG	872
DB	261	AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly	280
QY	873	GAAGAAGCGTATCATCCCTTTTGTATCGCAGTCACTTGAACCTGTTCCACATCAATACGTT	932
DB	281	GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal	300
QY	933	GATGTCGATACCAACGAATTTGTAAGAAGTGAGCAGCTCTTAACAGCTAGCGAACGTAAAC	992
DB	301	AspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn	320
QY	993	TTAGACTTCAGAGATTTATACGATCTCTCGTGATAAGGCTAAACTACTCTACAAATCTC	1052
DB	321	LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu	340
QY	1053	GATGCTTTTGGTATTATGGACTATACCTTAACCTGGAAGCTAGAGGATATACGATGAC	1112
DB	341	AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp	360
QY	1113	ACCAACGCTATCATACCCGTTTATATGCGCAAGCGACCCCAAGGAGAGAATGCTAGCTAC	1172
DB	361	ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr	380
QY	1173	CATTAGCT 1181	
DB	381	HisLeuAla 383	
RESULT 2			
AA	10194	AAAR10194 standard; protein; 414 AA.	
XX	AC	AAAR10194;	
XX	DT	28-MAR-1991 (first entry)	
XX	DE	Streptokinase encoded by synthetic gene.	
XX	KW	streptokinase; thrombolytic agent; myocardial infarction.	
XX	OS	Synthetic.	
XX	PN	EP407942-A.	
XX	PD	16-JAN-1991.	
XX	PF	11-JUL-1989; 89JP-00179432.	
XX	PR	11-JUL-1989; 89JP-00179432.	
XX	PR	27-NOV-1989; 89JP-00307957.	
XX	PR	11-APR-1990; 90JP-00096830.	
XX	PA	(SAKA) OTSUKA PHARM FACTOR.	
XX	PI	Majima E, Ogino K, Ono K, Sakata Y, Uenoyama T;	
XX	DR	WPI; 1991-016179/03.	
XX	DR	N-PSDB; AAQ10230.	
XX	PT	Synthetic gene encoding streptokinase - scale, high purity prodn. of	
XX	PT	streptokinase used as a thrombolytic agent.	
XX	PS	Claim 1; Page 51; 76pp; English.	

CC Streptokinase and its derivatives can be produced in large quantities
CC with high purity for use as thrombolytic agents in patients with lung
CC thrombus or myocardial infarction. See also AAR10195-R10200

XX Sequence 414 AA;

Alignment Scores:

Pred. No.:	1,13e-178	Length:	414
Score:	1984.00	Matches:	383
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	73.05%	Indels:	0
Ds:	2	Gaps:	0

US-09-940-235-9 (1-1541) x AAR10194 (1-414)

```

QY 33 AATAGCTGGCTCTGAAGGCTACTAGATCGTCTTCTGTAAATTAACAGCCAAATGGTTGT 92
DB 1 IleaIaGlyProGIuIrrPleuLeuAAspArgProSeValAAsnAsnSerGIuLeuVal 20
QY 93 AGCGTTGCTGTAAGCTGTGAGGAGCAATCAAGACATTAGTCTTAAATTTTGAATC 152
DB 21 SerValAaGIyThrValGIuGIyTrnAsnGlnAspLLeuSeruLysPhePheGIuIle 40
QY 153 GATTAACATCAAGACCTGCTCATGAGAGAAAGACAGAGAGAGCTTAACTCAAAATCA 212
DB 41 AspLeuThrSerArgProAlaHisGIyGIyLeuThrGIuGIuGIuLysSerProLysSer 60
QY 213 AAACCATTTGCTACTGATAGTGCGCGATGTCACATTAATTTGAAGAAAGCTGACTTA 272
DB 61 LysProPheAlaThrAspSerGIyAlaMetSerHisLysLeuGIuLysAlaAspLeu 80
QY 273 AAGGCTATCAAGAAATTAATGATCGTAACTGTCACAGTAAACAGCACTACTTGAGTGC 332
DB 81 LysAlaIleGIuGIuGIuLeuIleAlaAsnValHisSerAsnAspArgLyrPheGIuVal 100
QY 333 ATTGATTTTGAAGCGATCAACCATTAAGTCAAGTCAAGAGCTCTTGTGTCAG 392
DB 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGIyValLyrPheAlaAsp 120
QY 393 AAAGATGGTTCGGTAACTTGGCGAGCCAACTGTCCAAGAAATTTTGTAGCGGACAT 452
DB 121 LysAspGIySerValThrLeuProThrGlnProValGIuGIuLysLeuSerGIyHis 140
QY 453 GTGCGCGTGAACCATTAAGAAACCAATCAAAACCAACAGCAATTTGTGATGTG 512
DB 141 ValArgValArgProLyrGIyGIuLysProIleGIuAlaAsnGlnAlaLysSerValAspVal 160
QY 513 GAATATTAAGTACAGTTTACTCCCTTAAACCTGATGAGATTTTGACAGCGGTCAAA 572
DB 161 GIuLyrTrnValGlnPheThrProLeuAsnProAspAspPheArgProGIyLeuLys 180
QY 573 GATTAAGCTATTGAAAACTAGCTATCGGTGACACCATCAATCTCAAGAAATTA 632
DB 181 AspThrLysLeuLeuLysThrIleuAlaIleGIyAspThrIleThrSerGIuGIuLeu 200
QY 633 GCTCAAGCAAAAGCAATTTTAAACAAAAACCAACCAAGGCTTATACGATTATGAACGTGAC 692
DB 201 AlaGlnAlaGlnSerLLeuLeuAsnLysAsnHisProGIyLyrThrIleLyrGIuArgAsp 220
QY 693 TCCTCAATCGTCAATCAATGACATTTTCCGTAACGATTTTACCAATGATCAAGAG 752
DB 221 SerSerIleValThrHisAspAsnAspLLeuPheArgThrIleLeuProMetAspGIuGIu 240
QY 753 TTTACTTACCGTGTAAATAATCGGGAAACAGCTTATAGATCAATTAATAAAATCGGCTG 812
DB 241 PheThrLyrArgValLysAsnAsnArgGIuGIuGlnAlaLyrArgLLeuAsnLysLysSerGIyLeu 260
QY 813 AATGAAGAAATTAACCAACATGACCTGATCTGTAGAAATTAATTCGCTTAAAAAAAGG 872
DB 261 AsnGIuGIuLLeuAlaAsnAsnThrAspLeuLLeuSerGIuLysLyrLyrValLeuLysGIy 280
QY 873 GAAAAGCCGTATGATCCCTTGTGATGCGAGTCACTTGAACCTGTTCAACCATCAATACGTT 932

```

```

DB 281 GIuLysProLyrAspProPheAspArgSerHisLysLysLeuPheThrIleLysLyrVal 300
QY 933 GATGTCATATCCCAACGAATTCCTTAAAGAGAGAGCTTAAACGTAAGGAACGTAC 992
DB 301 AspValAspThrAsnGlnLeuLeuLysSerGIuGIuLeuThrAlaSerGIuArgAsn 320
QY 993 TTAGACTTCAGAGATTTATACGATCTCGTGAATAGGCTTAAACTTCAACAACTGC 1052
DB 321 LeuAspPheArgAspLeuLyrAspProArgAspLysAlaLysLeuLyrAsnAsnLeu 340
QY 1053 GATGCTTTGGTATTAATGACATAACCTTAACTGAAAGAGTAAGAGATTAACAGATGAC 1112
DB 341 AspAlaPheGIyLLeuMetAspLyrThrLeuThrGIyLysValGIuAspAsnHisAspAsp 360
QY 1113 ACCAAGCGTATCATACCGTTTATATAGGCAAGCGCAAGCAAGAGAGATGTAAGTAC 1172
DB 361 ThrAsnArgLLeuIleThrValLyrMetGIyLysArgProGIuGIuAsnAlaSerLyr 380
QY 1173 CATTAGCT 1181
DB 381 HisLeuAla 383

RESULT 3
AAR63120
ID AAR63120 standard; protein; 414 AA.
XX
AC AAR63120;
XX
DT 16-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 16-NOV-1994 (first entry)
XX
DE Streptokinase.
XX
KW Streptokinase; SK; Streptococcus equisimilis; plasminogen;
KM myocardial infarction.
XX
OS Streptococcus dysgalactiae subsp. equisimilis.
XX
XX
XX Key Location/Qualifiers
XX FT Region 1..352
XX FT /note= "claim 3, see CC"
XX FT Region 14..414
XX FT /note= "claim 1, see CC"
XX FT Region 120..352
XX FT /note= "claim 3, see CC"
XX FT Region 244..414
XX FT /note= "claim 3, see CC"
XX FT Region 244..352
XX FT /note= "claim 2, see CC"
XX
XX MO9407992-A1.
XX
XX PD 14-APR-1994.
XX
XX 05-OCT-1993; 93WO-US009502.
XX
XX 05-OCT-1992; 92US-00956692.
XX
XX 29-SEP-1993; 93US-00128299.
XX
XX (GENO ) GEN HOSPITAL CORP.
XX
XX (HARD ) HARVARD COLLEGE.
XX
XX PI Reed GL;
XX
XX DR WPI, 1994-135561/16.
XX
XX DNA encoding a polypeptide which binds to plasminogen and corresponds to
XX PT region of streptokinase - useful to detect plasminogen in a sample and to
XX PT treat myocardial infarction.
XX
XX PS Disclosure; Page 40-41; 62pp; English.

```

XX Nucleic acid comprising a sequence encoding amino acids 14-414 of
CC streptokinase, which binds to plasminogen and does not have a sequence
CC comprising amino acids 60-414 is new. The polypeptide pref. comprises
CC amino acids 244-352, 1-352, 120-352 or 244-414. (Updated on 25-MAR-2003
CC to correct PN field.) (Updated on 16-OCT-2003 to standardise OS field)
XX
SQ Sequence 414 AA;
Alignment Scores:
Pred. No.: 1.13e-178 Length: 414
Score: 1984.00 Matches: 383
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 73.05% Indels: 0
DB: 2 Gaps: 0
US-09-940-235-9 (1-1541) x AAR63120 (1-414)
QY 33 ATAGCTGGTCTGCTGCTAGTCTCTCTGTAATAACAGCAATTTGGTTGTT 92
DB 1 IleaAGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
QY 93 AGCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 152
DB 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 40
QY 153 GATCTAACATCAGCAGCTGCTCATGGAGGAAGAGAGAGAGAGAGAGAGAGAG 212
DB 41 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 60
QY 213 AAAACATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 272
DB 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeu 80
QY 273 AAGGCTATTCAAGAACATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 332
DB 81 LysAlaIleGlnGluLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100
QY 333 ATTGATTTTGCAGGAGTCAACCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 392
DB 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
QY 393 AAAAGATGTTGGTACCTTGGTCCGACCCCACTGCTCAAGAAATTTTGGCTGCTGCTGCT 452
DB 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
QY 453 GTGCGCGTTAGACCATATAAGAAAAACCAATACAAAAACCAAGCGAAATCTGTTGATGTG 512
DB 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
QY 513 GAATATCTGTACGTTACTCCCTTAAACCTGTATGACGATTTTCAGACAGGTCTCAAA 572
DB 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
QY 573 GATATACGCTATTCAACACATAGCTATCGGTGACACCATCACATCTCAAGNATTACTA 632
DB 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 200
QY 633 GCTCAAGCACAAGCATTTTAAACAAAAACCAACCCAGGCTATACGATTTTATGACGTGAC 692
DB 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 220
QY 693 TCCTCAATGCTCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 752
DB 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
QY 753 TTTACTTACCGTGTAAAAATCGGACAGAGCTTATAGATCAATATAAAAAATCTGCTCG 812
DB 241 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 260
QY 813 AATGAAGAAATAAACAACATGACCTGATCTCTGAGAAATATTACCTCTCTTAAAAAGGG 872

DB 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280
QY 873 GAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCAACATCAATACGTT 932
DB 281 GluLysProTyrAspPropheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
QY 933 GATGCTGATACCAACGAATTTGCTAAAAAGTGCAGCGCTCTTAAACAGCTAGCGAACGTAAC 992
DB 301 AspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320
QY 993 TTAGACTTCAGAGATTTATACGATCTCTGCTGATAAGGCTAACTACTCTCAACAATCTC 1052
DB 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
QY 1053 GATGCTTTTGGTATTATGAGTACTTAACTGGAAGAGTACGAGGATATACGATGAC 1112
DB 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
QY 1113 ACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAGAGAGAGATGCTAGCTAC 1172
DB 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 380
QY 1173 CATTTAGCT 1181
DB 381 HisLeuAla 383
RESULT 4
AA24794
ID AAY24794 standard; protein; 414 AA.
AC AAY24794;
XX
DT 17-OCT-2003 (revised)
DT 26-AUG-1999 (first entry)
XX
DE Streptococcus equisimilis native streptokinase.
XX
KW Streptococcus; streptokinase; fibrin-dependent plasminogen activator;
KW nSK; rSK; bacterial; blood clot; thrombotic condition;
KW myocardial infarction; venous thrombosis; pulmonary embolism;
KW cerebral thrombosis; graft thrombosis; arterial thrombosis.
XX
OS Streptococcus dysgalactiae subsp. equisimilis.
XX
PN WO9931247-A1.
XX
PD 24-JUN-1999.
XX
PF 15-DEC-1998; 98WO-US026694.
XX
PR 15-DEC-1997; 97US-0069497P.
XX
PA (HARD) HARVARD COLLEGE.
XX
PI Reed GL;
XX
DR WPI; 1999-395183/33.
DR N-PSDB; AAX80492.
XX
PT N-terminally deleted streptokinase.
XX
PS Claim 30; Page 60-61; 73pp; English.
XX
CC The present invention describes an isolated bacterial protein that
CC induces fibrin-dependent plasminogen activation in a pharmaceutical
CC composition for dissolving blood clots. Also described are: (1) a
CC composition comprising an isolated modified streptokinase, the
CC modification being removal of amino acid residues in the amino terminus;
CC (2) a method for dissolving a blood clot in a subject, comprising
CC administering to the subject a fibrin-dependent streptokinase protein; a
CC nucleic acid (1) encoding a modified bacterial streptokinase; (3) an
CC expression vector comprising (1); and (4) a host cell transformed with
CC the expression vector of (3). The pharmaceutical composition comprising a

CC bacterial fibrin-dependent plasminogen activator is useful for dissolving
 CC blood clots in patients with a thrombotic condition, e.g. myocardial
 CC infarction, venous thrombosis, pulmonary embolism, cerebral thrombosis,
 CC graft thrombosis and arterial thrombosis. The modified streptokinase can
 CC also be used in non-human mammals. Streptokinase activation of
 CC plasminogen is at least 10-fold, preferably 100-fold greater in the
 CC presence of fibrin than in the absence of fibrin. The modified
 CC streptokinase has at least one amino acid substitution that inactivates a
 CC substrate site for proteolytic cleavage. This reduces the rate of
 CC degradation of the streptokinase at least two-fold. The present sequence
 CC represents native streptokinase (nsk). (Updated on 17-OCT-2003 to
 CC standardise OS field)

XX Sequence 414 AA;

Alignment Scores:

Pred. No.:	1,13e-178	Length:	414
Score:	1984.00	Matches:	383
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	23.05%	Indels:	0
DB:	2	Gaps:	0

US-09-940-235-9 (1-1541) x AAY24794 (1-414)

```

QY 33 ATAGCTGCTCTGAAATGGCTACTAGATCGCTTCTGTAAATAACAGCAATTGGTGT 92
DB 1 IleaIaGIYProGIuITrIeuLeuAaPArGProSeValaIaSnHnSerGIuLeuVal 20
QY 93 AGCGTGTCTGACTGTGTAGAGGAGCAATCAAGACTTAAGTCTTAATTTTGAATC 152
DB 21 SerValaIaGIYTrValaGIuGIYTrAaSnGIaSPrIeSerIeuIySPhePheGIuIle 40
QY 153 GATCTAACATCAACGCTCTCATGAGAGAAACAGAGAGAGCTTAAGTCCAAATCA 212
DB 41 AspleuTrIeSerArProAlaHISGIYGIYLeThGIuGIuGIYLeuSerProIySer 60
QY 213 AAACATTTGTCTACTGATAGTGGCGGATGTCATTAATTTTGAAGCTTACTA 272
DB 61 LysPProPheAlaIeThAsPserGIYAlaMetSerHISLeuGIuIySalAsPleuLeu 80
QY 273 AAGGCTATTCAAGAAACAATTGATCGCTAACGTCACAGTACAGCACTTCTTGAGTGC 332
DB 81 LysAlaIaIeGIuGIuGIuLeuIleAlaSnValaHISerAsnAsPArGIYrPheGIuVal 100
QY 333 ATTGATTTTGCAGGATGCAACCAATTACTGATGGAACGGCAAGTCTTACTTGCTGAC 392
DB 101 IleAsPpHeAlaSerAsPAlaThTrIeThAsPArGAsnGIYLeValITyrPheAlaAsP 120
QY 393 AAAGATGTTGGTGTAACTTGGCGAGCAACCTGTCACAGAAATTTTGTAAACGGACAT 452
DB 121 LysAsPGIYSerValThIeuProInGIuProValGIuGIuIeLeuSerGIYHIS 140
QY 453 GTGCGCGTTAGACCATATTAAGAAAAACAATCAAAACAAGCAAGAAATCTGTGATGTC 512
DB 141 ValaGValaGProGIYrIySGIuYSProIleGIuHnSnGIuAlaIySSerValaAsPVal 160
QY 513 GAATTAAGCTTAAGTTAAGTTAAGTTAAGTTAAGTTAAGTTAAGTTAAGTTAAGTTAAG 572
DB 161 GIuYrThIeValaGIuPheThIeProLeuAaPArGAsPArGpHeArGIYLeuIyS 180
QY 573 GATATAAGCTTAAGTTAAGTTAAGTTAAGTTAAGTTAAGTTAAGTTAAGTTAAGTTAAG 632
DB 181 AsPTrIySLeuLeuYThIeValaIeGIYAsPTrIeThIeSerGIuIeLeuLeu 200
QY 633 GCTCAAGCAAAAGCATTTTAAACAAAAACCAACCGAGGATTCGATTAAGAAAGTGCAC 692
DB 201 AlaGIuAlaGIuSerIleuSnIySAsnHISProGIYrThIeYrGIuIaGAsP 220
QY 693 TCCTCAATCGTCACTCATGACATGACATTTTCCGATGATTTTCAATGATGATGATGATG 752
DB 221 SerSerIleValaThHISAsPAsnAsPArIePheArGIYThIeLeuProMetAsPGIuGIu 240

```

```

QY 753 TTACTTACCGTGTAAATAATCGGAAACAAGCTTATAGATCAATCAATAAATCGTCTG 812
DB 241 PheThIeYrArGValaIySAsnAArgGIuInIaIeYrArGIleAsnIySLeuSerGIYLeu 260
QY 813 AATGAAGAAATTAACAACACTGACCTGATCTCTGAGAAATATTAAGTCTTAAAGAGG 872
DB 261 AsnGIuGIuIeAaAsnThIeAsPleuIeSerGIuIyYrYrValleuIySGLY 280
QY 873 GAAAGCGGTATGATCCCTTGGATCGGATGCACTTGAACGTTGACCAATCAATACCTT 932
DB 281 GIuYSPProIyArSPProPheAsPArGSerHISLeuIySLeuPheThIeIySTrVal 300
QY 933 GATGTCAATCAACAAGAAATTCCTTAAGAGAGAGCTCTTAACAGTACGAAAGTAAAC 992
DB 301 AsPValaSPTrIeAaSnGIuLeuLeuIySerGIuGIuLeuThIeAlaSerGIuArGAsn 320
QY 993 TTAAGCTTCAGAGATTTATACATCTCTGTATTAAGCTTAAAGCTTAACTTACAAATCTC 1052
DB 321 LeuAsPpHeArGAsPleuYrAsPProArGAsPArGIYSalalYSLeuLeuYrAsnAsnLeu 340
QY 1053 GATGCTTTGGTATTATGACTATACCTTAAGTGAAGAGTAAAGATTAACAGATGAC 1112
DB 341 AsPAlaPheGIYIleMetAsPTrIeThIeThrGIYValaGIuAsPAsnHISAsPAsP 360
QY 1113 ACCAAGCTATCATACCGTTATATGAGCAAGCAACCGAAGAGAGATGCTAGCTAC 1172
DB 361 ThrAsnArGIleIeThIeThIeValaIeMetGIYAsArGProGIuGIuIeAsnHISerYr 380
QY 1173 CATTAGCT 1181
DB 381 HisIeuAla 383

```

RESULT 5
 AAW94664
 ID AAW94664 standard; protein; 414 AA.
 AC AAW94664;
 XX
 DT 17-OCT-2003 (revised)
 DT 04-MAY-1999 (first entry)
 XX
 DE Streptococcus equisimilis native streptokinase.
 XX
 KW Streptococcus equisimilis H46A; streptokinase; mutant; fibrinolytic;
 KW plasma clot; hydrolysis; haemolytic Streptococcus; plasminogen; plasmin;
 KW serine protease; fibrin; blood clot; thrombolytic;
 KW vascular thromboembolytic symptom; acute myocardial infarction;
 KW fibrinolysis; resistance.
 XX
 OS Streptococcus dysgalactiae subsp. equisimilis.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 14
 FT /note= "encoded by ACC"
 PN US5876999-A.
 XX
 PD 02-MAR-1999.
 XX
 PF 06-DEC-1995; 95US-00568393.
 XX
 PR 06-DEC-1995; 95US-00568393.
 XX
 PA (NASC-) NAT SCI COUNCIL.
 PI
 PI Wu H;
 XX
 DR WPI; 1999-189643/16.
 DR N-PSDB; AAX16632.
 XX
 PT Mutant streptokinase polypeptide - useful as plasmin-resistant
 PT thrombolytic agent.

PS Claim 1; Col 7-10; 17pp; English.

XX The present invention describes a mutant streptokinase (SK) polypeptide in which at least one of the amino acids in the Pro58-Lys59-Ser60-Lys61 segment of the corresponding native SK is replaced by another amino acid. The present sequence represents native SK. SK is a secretory protein of haemolytic Streptococcus able to activate human plasminogen (HPlg) to plasmin (HPlm), which is a serine protease able to catalyse the hydrolysis of fibrin in blood clots. The SK is useful as a thrombolytic agent in the treatment of vascular thromboembolytic symptoms such as acute myocardial infarction. Compared with wild-type SK, the K59E mutant is more resistant to degradation by human plasmin and is more effective both in acting as a fibrinolytic agent and in activating human plasminogen. (Updated on 17-OCT-2003 to standardise OS field)

XX Sequence 414 AA;

Alignment Scores:

Pred. No.:	1.13e-178	Length:	414
Score:	1984.00	Matches:	383
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	73.05%	Indels:	0
DB:	2	Gaps:	0

US-09-940-235-9 (1-1541) x AAW94664 (1-414)

QY 33 ATAGCTGCTCGTACTAGTCTCTCTGTAATAACAGCAATGGTGT 92

Db 1 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20

QY 93 AGCGTTGCTGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTGAAATC 152

Db 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPheGluIle 40

QY 153 GATCTAATCATCAGCACTCTCATGGAGGAAAGACAGCAAGGCTTAAGTCCAAATCA 212

Db 41 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 60

QY 213 AAACATTGCTACTGATAGTGGCGGATGTCACATAACTTGAGAAAGCTGACTACTA 272

Db 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 80

QY 273 AAGGCTATTCAAGAACAAATTGATCGCTACGTCACAGTAACGACGACTACTTTGAGTCC 332

Db 81 LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100

QY 333 ATTGATTTTGAAGCGATGCAACCATTTACTGATCGAAACGGCAAGGCTCTACTTTGCTGAC 392

Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120

QY 393 AAAGATGTTGCTGTAACCTTCCGACCCCACTCTGTCACAGAAATTTTGTAAAGCGACAT 452

Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140

QY 453 GTGCGGTTAGCAACCATATAAGAAACCAATACAAACCAACCGAAATCTGTGATGTCG 512

Db 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160

QY 513 GAATATATCTGACGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAAA 572

Db 161 GlutyrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuLys 180

QY 573 GATACTAGCTATTGAAACACTAGCTATCGGTGACACCATACATCTCAAGAAATPACTA 632

Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 200

QY 633 GCTCAAGCACAAAGCATTTTAAACAAACCAACCCAGGCTATACGTTTATGACGTCGAC 692

Db 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 220

QY 693 TCCTCAATCGTCATGACAAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 752

Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240

QY 753 TTTACTTTACCGTGTATAAAATCGGAAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 812

Db 241 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 260

QY 813 AATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAGGG 872

Db 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280

QY 873 GAAAGCCGATGATCCCTTTGATCGCAGTCACCTTGAACCTGTTACCAATCAAAATAGCTT 932

Db 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300

QY 933 GATGTCGATACCAACGAATTTCTAAAAGTGACGAGCTTTAAACAGCTAGCGAACGTAAC 992

Db 301 AspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320

QY 993 TTAGACTTCAGAGATTTTATACGATCCTCGTGATAAGGCTAAAGCTACTCTACAACAATCTC 1052

Db 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340

QY 1053 GATGCTTTGCTATTATGGAATACCTTAACCTGGAAGAGTAGAGGATAATCAGCATGAC 1112

Db 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360

QY 1113 ACCAACGATATCAACCGTTTATATGCGCAAGCGACCCGAGAGAGAGATGCTAGCTAC 1172

Db 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 380

QY 1173 CATTTAGCT 1181

Db 381 HisLeuAla 383

RESULT 6

AA01556

ID AAY01556 standard; peptide; 414 AA.

XX AAY01556;

XX 17-OCT-2003 (revised)

DT 18-JUN-1999 (first entry)

XX Native streptokinase protein sequence.

DE Antigenic peptide; streptokinase; streptokinase-specific antibody;

XX Thrombolytic activity; thrombolytic therapy; glomerulonephritis;

KW rheumatic fever.

XX Streptococcus dysgalactiae subsp. equisimilis.

OS WO9908698-A1.

PN 25-FEB-1999.

XX 18-AUG-1998; 98WO-US017114.

PF 18-AUG-1997; 97US-0055911P.

PR (HARD) HARVARD COLLEGE.

PA (GEHO) GEN HOSPITAL CORP.

XX Reed GL, Parhami-Seren B;

PI WPI; 1999-190113/16.

XX New polypeptides which bind streptokinase-specific antibodies - useful in thrombolytic therapy.

PT Disclosure; Page 12; 44pp; English.

XX The present sequence represents a native streptokinase. The specification describes a polypeptide which binds to a streptokinase-specific antibody

CC and prevents the antibody binding to native streptokinase. The CC specification also describes a synthetic polypeptide (PI) comprising an epitope which binds to an streptokinase-specific antibody and reduces CC thrombolytic activity of streptokinase. PI is used in thrombolytic CC therapy and to prevent or treat glomerulonephritis and rheumatic fever. CC (Updated on 17-Oct-2003 to standardise OS field)

XX Sequence 414 AA:

Alignment Scores:

Pred. No.:	1 13e-178	Length:	414
Score:	1984.00	Matches:	383
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	73.05%	Indels:	0
DB:	2	Gaps:	0

US-09-940-235-9 (1-1541) x AA01556 (1-414)

```

OY 33 ATAGCTGCTCCTGAATGGCTACTAGATCGTCTTCTGTAATAACAGCAATGGTGT 92
DB 1 TlealaglyProglutirpleuleAaspArgProSerValashnenserGlnleuVal 20
OY 93 AGCGTTGCTGTACTGTGAGGAGCAATCAAGACATTAGCTTAAATTTTGAATC 152
DB 21 SerValalaglyThrValGluGlyThrAsnGlnaspIleSerleuyspPhepGluIle 40
OY 153 GATTTAAACATCAGACCTGCTCATGAGGAAAGACAGACGAGCTTAAGTCCAAATCA 212
DB 41 AspLeuThrSerArgProalashGlyGlyLysThrGlnGlnGlyLeuSerProLysSer 60
OY 213 AAACATTGCTACTAGTAGTGGCGGATGTCATTAACCTTGAGAAAGCTTACTA 272
DB 61 LysProPhealathraspserGlyalameSerHisLysleuGluysalAaspLeu 80
OY 273 AAGGCTATTCAAGAACATTTGATCGCTAACGTCACAGTACGACGACTCTTGAGATC 332
DB 81 LysalalieglnGlnleuIlealashenValHisSerashnaspArgyrPheGluVal 100
OY 333 ATTTGATTTTGCAGGACGATGACCATTTACTGATCGAAACGGCAAGCTTACTTGCTGAC 392
DB 101 IleaspPhealaseraspalathrIleThrAspArgasnGlyysValTyrPhealAasp 120
OY 393 AAAGATGTTGGCTTAACCTTGGCGGACCCGACCTGTCACAGATTTTGTAAACGGGACAT 452
DB 121 LysaspGlySerValThrleuProthGlnProValGlnGlnPheleuSerGlyHis 140
OY 453 GTGCGGTTAGACCATATTAAGAAAAACAATACAAACCAAGCGAAATCTGTTGATGTC 512
DB 141 ValArgValArgProIlyrlysgIuysProIleGlnashnGlnalalysserValaspVal 160
OY 513 GAATATCTGTACAGTTTACTCCCTTAAACCTGATGACATTTTACAGCAGGCTTCAAA 572
DB 161 GluYrThrValGlnPheThrProleuAenProaspAspPheargProGlyLeuLys 180
OY 573 GATTACTTAAGCTATTTGAAAACTAGCTATGCGTGACACCATCTCAAGATTTACTA 632
DB 181 AspThrLysleuLeuysrThrleuValalieglyAspThrIleThrserGlnGluLeu 200
OY 633 GGTCAAGCACAAGCATTTTAAACAAAAACCAACCGGCGGTATTCGATTTATGAACGGAC 692
DB 201 AlaglnalaglnSerIleleuashnlyshashaspGlyrYrThrIleIyrgluashp 220
OY 693 TCCTCAATGTCATCTAGTACATGACATTTTCCGTACGATTTTACCAATGATCAAG 752
DB 221 SerSerIleValThrHisaspashnaspIlePheargThrIleleuProketaaspGlnGlu 240
OY 753 TTTACTTACCGTGTAAAAATCGGGAACAAGCTTATAGATCAATAAAAATCTGGCTG 812
DB 241 PheThrYrArgVallyshasnargIuGlnalalyrArgIleashnlysserGlyLeu 260
OY 813 AATGAAGAAATAACAACACTGACCTGATCTCGAAGAAATATATACGCTTAAAAAAGG 872

```

```

DB 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGlyLysTyrValIleuLysGly 280
OY 873 GAAAGCGGTATGATCCCTTTGATCGCAGTCACTGAAACGTTTACCATCAATACGTT 932
DB 281 GluyspProIlyrAspProPheaspArgSerHisleuLysleuPheThrIleYsrVal 300
OY 933 GATGTCGATCAACAAGAAATGCTTAAAAAGTGAGCAGCTTTTACAGCTAGCAACGTAC 992
DB 301 AspValaspThrAsnGlnleuLeuLysSerGlnGlnleuThrAlaSerGluashn 320
OY 993 TTAGACTTCAGAGATTTATACGATCTGTGATTAAGGCTAAACCTTACCAACATCTC 1052
DB 321 LeuaspPheargaspLeuLysAspProargaspLysalalySleuLeuLysAsnashn 340
OY 1053 GATGCTTTTGCTATTTATGACTATACCTTAACTGGAAGAGTACAGATATCACATGAC 1112
DB 341 AspAlapheGlyIleMetaspLysrThrleuThrGlyLysValGluaspashnIleasp 360
OY 1113 ACCAAGCTATCATACCGTTTATATGAGGCAAGGACCGAAGAGAGAAATGCTAGCTAC 1172
DB 361 ThrashnArgIleIleThrValTyrMetGlyLysargProGluGlyGluashnAlaSer 380
OY 1173 CATTAGCT 1181
DB 381 HisleuAla 383

```

RESULT 7
AA090282
ID AA090282 standard; protein; 414 AA.
XX
AC AA090282;
XX
DT 12-SEP-2003 (revised)
DT 13-OCT-2000 (first entry)
XX
DE S. equisimilis streptokinase.
XX
KW Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
KW plasminogen; human; fibrinectin; thrombolytic therapy;
KW cardiovascular disorder; fibrinectin.
XX
OS Streptococcus dysgalactiae subsp. equisimilis.
XX
EN EPI024192-A2.
XX
PD 02-AUG-2000.
XX
PF 23-DEC-1999; 99EP-00310541.
XX
PR 24-DEC-1998; 98IN-DE003825.
XX
PA (COUL) CSIR COUNCIL SCI IND RES.
XX
PI Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
XX Yadav M;
XX
DR WPI; 2000-516032/47.
DR N-PSDB; AAA37633.
XX
PT Hybrid streptokinase-fibrin binding domain polypeptides useful for
PT thrombolytic therapy comprises a streptokinase fused with fibrin binding
PT domains of human fibrinectin.
XX
PS Example 3; Fig 3; 58pp; English.
XX
CC This sequence represents the human Streptococcus equisimilis
CC streptokinase protein sequence. The invention relates to a hybrid
CC plasminogen activator (PA) comprising a polypeptide fusion between
CC streptokinase (SK), which are capable of plasminogen (PG) activation, and
CC fibrin binding regions of human fibrinectin, which are from fibrin
CC binding domains (PBD) 4 and 5 or 1 and 2. The hybrid PA possesses the
CC ability to bind with fibrin independently and also characteristically
CC retains a PG activation ability which becomes evident only after a

CC pronounced duration, or lag, after exposure of the PA to a suitable
 CC animal or human PG. The hybrid streptokinase-fibrin binding domain
 CC polypeptides are useful in thrombolytic therapy for various kinds of
 CC cardiovascular disorders. The hybrids have enhanced fibrin selectivity as
 CC well as kinetics of plasminogen activation that are distinct from that of
 CC natural streptokinase in being characterised by a temporary delay, or lag
 CC of several minutes in the natural rate of the catalytic conversion of
 CC plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins
 CC can bind tightly with fibrin in blood clots soon after introduction into
 CC the vascular system without significantly activating the circulating
 CC blood plasminogen to plasmin, thus aiding in the localisation of the
 CC plasminogen activation process to the site of pathological thrombus. This
 CC overcomes systemic plasminogen activation encountered during clinical use
 CC of streptokinase. (Updated on 12-SEP-2003 to standardise OS field)
 XX
 SQ Sequence 414 AA;

Alignment Scores:
 Pred. No.: 1,13e-178 Length: 414
 Score: 1984.00 Matches: 383
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 73.05% Indels: 0
 DB: 3 Gaps: 0

US-09-940-235-9 (1-1541) x AAY90282 (1-414)

QY 33 ATAGCTGCTCTGAATGGCTACTAGATCGCTCTCTGTAATAACAGCCAAATGGTGT 92
 DB 1 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnSerGlnLeuValVal 20
 QY 93 AGCTTCTGCTACTGTTGAGGGGACGATCAAGACATTAGTCTTAATTTTGAATC 152
 DB 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPheGluIle 40
 QY 153 GATCTAATCAGCACTGCTCATCGAGGAAAGACAGAGCAAGCTTAAGTCCAAATCA 212
 DB 41 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 60
 QY 213 AAACATTTGCTACTGATAGTGGCGGATGTCACATAAATCTGAGAAAGCTCACTACTA 272
 DB 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 80
 QY 273 AAGCTATTCAAGACAAATGATCGCTAACGTCACAGTCAAGCACTACTTGGTGC 332
 DB 81 LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100
 QY 333 ATTGATTTTCAAGCGATGCAACCAATTACTGATCGAAACGCAAGTCTACTTTGTGAC 392
 DB 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
 QY 393 AAAGATGTTCCGTTAACTTGCACGCCCAACCTGTCCAGAAATTTTGTAAAGCGACAT 452
 DB 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
 QY 453 GTGGCGTTAGACCATATAAGAAAAACCAATPACAAAACCAAGCGAAATCTGTTGATGG 512
 DB 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
 QY 513 GAATATATCTGATGATTTTACTCCCTTAAACCTGATGAGTTCAGACCAAGTCTCAAA 572
 DB 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
 QY 573 GATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTTACTA 632
 DB 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 200
 QY 633 GCTCAGCACAAAGCAATTTTAAACAAAACCAACCCAGGCTATACGATTTATGAACTGAC 692
 DB 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 220
 QY 693 TCCTCAATCGTCACTCAGCAATCACATTTTCCGTACGATTTTACCAATGATCAGAG 752

DB 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
 QY 753 TTTACTACCGTCTTAAATCGGAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 812
 DB 241 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 260
 QY 813 AATGAAGAATAAACCAACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 872
 DB 261 AsnGluGluLeuAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280
 QY 873 GAAACCCGTATGATCCCTTTGATCGAGTCACCTGAAACTGTTCACCATCAATACGTT 932
 DB 281 GluLysProTyrAspProPheAspArgSerHisLeuLeuLeuPheThrIleLysTyrVal 300
 QY 933 GATGTCGATACCAACGAATTGCTTAAAGTGACAGCTCTTAACAGCTACGCAACGTAAC 992
 DB 301 AspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320
 QY 993 TTAGACTTCAGAGATTATACGATCTCTGATGTAAGCTAAATCTACTCTACACAATCTC 1052
 DB 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
 QY 1053 GATGCTTTTGGTATTATGACTATACCTTAACCTGGAAGTAGAGGATATACGATGAC 1112
 DB 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
 QY 1113 ACCAACCGTATCATACCGTTTATATGCGCAAGCGACCCGAGAGAGAGATGCTAGCTAC 1172
 DB 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 380
 QY 1173 CATTTAGCT 1181
 DB 381 HisLeuAla 383
 RESULT 8
 AAY24797
 ID AAY24797 standard; protein; 795 AA.
 XX
 AC AAY24797;
 XX
 DT 26-AUG-1999 (first entry)
 XX
 DE Streptokinase and maltose binding protein fusion protein.
 XX
 KW Streptococcus; streptokinase; fibrin-dependent plasminogen activator;
 KW rSK; rSK; bacterial; blood clot; thrombotic condition;
 KW myocardial infarction; venous thrombosis; pulmonary embolism;
 KW cerebral thrombosis; graft thrombosis; arterial thrombosis.
 XX
 OS Streptococcus dysgalactiae subsp. equisimilis.
 OS Synthetic.
 XX
 PN W09931247-A1.
 XX
 PD 24-JUN-1999.
 XX
 PF 15-DEC-1998; 98WO-US026694.
 XX
 PR 15-DEC-1997; 97US-0069497P.
 XX
 PA (HARD) HARVARD COLLEGE.
 XX
 PI Reed GL;
 XX
 XX WPI; 1999-395183/33.
 DR N-PSDB; AAX80497.
 XX
 PT N-terminally deleted streptokinase.
 XX
 PS Example; Page 48-51; 73pp; English.
 XX
 CC The present invention describes an isolated bacterial protein that
 CC induces fibrin-dependent plasminogen activation in a pharmaceutical

CC composition for dissolving blood clots. Also described are: (1) a
 CC composition comprising an isolated modified streptokinase, the
 CC modification being removal of amino acid residues in the amino terminus;
 CC (2) a method for dissolving a blood clot in a subject, comprising
 CC administering to the subject a fibrin-dependent streptokinase protein; a
 CC nucleic acid (1) encoding a modified bacterial streptokinase; (3) an
 CC expression vector comprising (1); and (4) a host cell transformed with
 CC the expression vector of (3). The pharmaceutical composition comprising a
 CC bacterial fibrin-dependent plasminogen activator is useful for dissolving
 CC blood clots in patients with a thrombotic condition, e.g. myocardial
 CC infarction, venous thrombosis, pulmonary embolism, cerebral thrombosis,
 CC graft thrombosis and arterial thrombosis. Streptokinase activation of
 CC plasminogen is at least 10-fold, preferably 100-fold greater in the
 CC presence of fibrin than in the absence of fibrin. The modified
 CC streptokinase has at least one amino acid substitution that inactivates a
 CC substrate site for proteolytic cleavage. This reduces the rate of
 CC degradation of the streptokinase at least two-fold. The present sequence
 CC represents a streptokinase and maltose binding protein fusion protein
 CC from an example of the present invention

CC
 XX Sequence 795, AA:

Alignment Scores:
 Pred. No.: 1,48e-178 Length: 795
 Score: 1984.00 Matches: 383
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 73.05% Indels: 0
 DB: 2 Gaps: 0

US-09-940-235-9 (1-1541) x AAY24797 (1-795)

QY 33 ATAGCTGCTCCTGAATGCGTACTGATGCTCTTCTGTAATTAACAGCAATTGTTGTT 92
 Db 382 IleaIaGlyProGluTrpLeuLeuAparGProSerValaAsnSerGlnLeuVal 401
 QY 93 AGCGTTGCGTACTGCTGAGGGGAAGATCAAGACATTAGCTTAATTTTGAATC 152
 Db 402 SerValaIaGlyThrValaGlnGlyTrpAsnGlnAparIleSerLeuYsPheGluIle 421
 QY 153 GATTTACATCAAGACCTGCTCATGAGAGAAAGACAGCAAGCTTAAGTCCAAATCA 212
 Db 422 AspleuThrSerArgProIaHisGlyGlyThrGlnGlnIleuSerProIySer 441
 QY 213 AAACCATTTGCTACTGATAGTGCGCGGATGTCACATTAACCTTGAGAAAGCTTACTA 272
 Db 442 LysProPheAlaThrAspSerGlyAlaMetSerHisIleLeuGlnIulysAlaAparLeu 461
 QY 273 AAGGCTATTCAGAGAACAAATTGATCGCTAACCTCCACAGTAACAGCACTTGTAGGTC 332
 Db 462 LysAlaIleGlnGlnIleuIleAlaValHisSerAsnAspAspTYrPheGluVal 481
 QY 333 ATTGATTTTGAAGCGATCAACCATTTACTGATCGAAACGGCAAGGCTACTTGTGCTAC 392
 Db 482 IleAspPheAlaSerAspAlaThrIleThrAspAsnGlyYsValTYrPheAlaAsp 501
 QY 393 AAAAGATGCTTCGGTAACTTGCAGCAACCTGTCCAAGATTTTGTACAGCAAT 452
 Db 502 LysAspGlySerValThrLeuProThrGlnProValGlnGlnIleuLeuSerGlyHis 531
 QY 453 GTGCGCGTTAGACCATATAAAGAAAAACAAATCAAAACCAAGCGAAATGTGTGATGTG 512
 Db 522 ValArgValaArgProIyTrpGlySerProIleGlnAsnGlnAlaIysSerValaAspVal 541
 QY 513 GAATATACCTGTCAGTTTACTCCCTTAACCCGATGAGATTTTGACACAGGCTCCAA 572
 Db 542 GluTYrThrValaGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuYs 581
 QY 573 GATATAAGCTATTGAAAAACATAGCTATCGGTGACACCATCACTCAAGATTAATCA 632
 Db 562 AspThrIysLeuLeuYsThrIleuAlaIleGlyAspThrIleThrSerGlnIleuLeu 581

QY 633 GCTCAAGCACAAGACATTTTAAACAAAAACCAAGGCTATAGCAATTTATGAAGTAC 692
 Db 582 AlaGlnIaGlnSerIleLeuAsnIysAsnHisProGlyTYrThrIleIleTYrGluArgAsp 601
 QY 693 TCTCAATCGTACTCATGACATGACATGACATTTTCCGTAGCATTTTACCAATGATCAAG 752
 Db 602 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnIu 621
 QY 753 TTTACTTACCGGTAAATAATCGGGAACAACCTTATAGGACATTAATAAATTCGTCTG 812
 Db 622 PheThrTYrArgValIysAsnArgGlnGlnAlaTYrArgIleAsnIysIysSerGlyLeu 641
 QY 813 AATGAAGAAATTAACAACACTGACCTGATCTGTGAAGATATTAAGTCTTAATAAAGG 872
 Db 642 AsnGlnGlnIleAsnAsnThrAspLeuIleSerGlnIySTYrTYrValIleuIyIyGly 661
 QY 873 GAAAAAGCGTATGATCCCTTTGATCGAGTCACTTGAACCTGTTACCATCAATACGTT 932
 Db 662 GluYsProTYrAspProPheAspArgSerHisLeuYsLeuPheThrIleIyTYrVal 681
 QY 933 GATGTCATACCAACGAATTGCTTAATAAGTAGAGAGCTTTAAACGCTAGCAAGTAC 992
 Db 682 AspValaAspThrAsnGlnLeuLeuIysSerGlnIleuLeuThrAlaSerGluArgAsn 701
 QY 993 TTAGACTTCAGAGATTTATACGATCCTCGTATAGGCTAAACTACTTACACAAATGTC 1052
 Db 702 LeuAspPheArgAspLeuTYrAspProArgAspYsAlaIySLeuLeuTYrAsnAsnLeu 721
 QY 1053 GATGCTTTGGTATTAATGACCTATACCTTAACTGGAAGAGTAGAGATTAACAGTAC 1112
 Db 722 AspAlaPheGlyIleMetAspTYrThrLeuThrGlyIySValGluAspAsnHisAspAsp 741
 QY 1113 ACCAAGCGTATCATACCGCTTTATATGGGCAAGGCAAGCAAGATGCTAGTAC 1172
 Db 742 ThrAsnArgIleIleThrValTYrMetGlyIySArgProGlnGlnIyGluAsnAlaSerTYr 761
 QY 1173 CATTTAGCT 1181
 Db 762 HisLeuAla 764

RESULT 9
 ID AAR11829
 XX AAR11829 standard; protein: 531 AA.
 AC AAR11829;
 XX
 DT 08-JUL-1991 (first entry)
 XX
 DE FB-FB-SK fusion conjugate.
 XX
 KW Fibrin-binding protein; fibrinolysis; intravascular thrombi; fibrinogen;
 KW streptokinase; fusion protein.
 XX
 OS Staphylococcus aureus.
 XX
 Key Location/Qualifiers
 FH Peptide 2..61
 FT Peptide /label= FB monomer
 FT Peptide 62..119
 FT Peptide /label= FB monomer
 FT Peptide 120..531
 FT Peptide /label= streptokinase
 PN US5011686-A.
 XX
 PD 30-APR-1991.
 XX
 PF 15-NOV-1989; 89US-00437769.
 XX
 PR 21-SEP-1987; 87US-00099242.
 XX
 PA (CREA-) CREATIVE BIOMOLEC.
 XX

PI Pang RHL;
 XX WPI; 1991-140198/19.
 DR N-PSDB; AAQ11651.
 XX
 PT Imparting injectable fibrinolytic agent - with affinity for intravascular
 XX thrombus, by linking agent to fibrin binding domain.
 XX Disclosure; Fig 5; 18pp; English.
 PS
 XX
 CC The conjugate comprises an FB-FB dimer linked to streptokinase The FB
 CC fragment has selective affinity for fibrin, low affinity for fibrinogen,
 CC and minimal immunogenicity, imparting thrombus- targeting capability.
 CC See also AAR11821 and AAR11828
 XX
 XX
 SQ Sequence 531 AA;
 Alignment Scores:
 Pred. No.: 2,41e-178 Length: 531
 Score: 1981.00 Matches: 382
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 99.74% Mismatches: 0
 Query Match: 72.94% Indels: 0
 DB: 2 Gaps: 0
 US-09-940-235-9 (1-1541) x AAR11829 (1-531)
 QY 33 ATAGCTGCTCCTGAATGGCTACTAGATCGCTCTCTGTAATAACAGCCAAATGGTGT 92
 Db 118 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnThrGlnLeuValVal 137
 QY 93 AGCGTGTCTGGTACTGTTGAGGGGCGAATCAAGACATTAGTCTTAAATTTTTGAAATC 152
 Db 138 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuIysPheGluIle 157
 QY 153 GATCTAACATCATCGACCTGCTCATCGAGGAGAAAGACAGCAAGGCTTAAGTCCAAATCA 212
 Db 158 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 177
 QY 213 AAACCATTTGCTAGTATGATGGCGGATGTACATAAACTTGAGAAAGCTGACTACTA 272
 Db 178 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 197
 QY 273 AAGGCTATTCAAGAACATTGATCGTACGTCACAGTAAACGACGACTACTTTGAGGTC 332
 Db 198 LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspTyrPheGluVal 217
 QY 333 ATTGATTTTGCAGCGATGCAACCACTTACTGATCGAAACGCGCAAGGCTCTACTTTGCTGAC 392
 Db 218 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 237
 QY 393 AAAGATGGTTGGTAACTTGGCGGACCCAACTGTCCAAAGAAATTTTGTAGCGGACAT 452
 Db 238 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 257
 QY 453 GTGCGGTTAGACCATATAAGAAACCAATACAAACCAAGCGAAATCTGTTGATG 512
 Db 258 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 277
 QY 513 GAATATACCTGTACAGTTTACTCCCTTAAACCTGTATGACGATTTTACAGCAGGTCTCAAA 572
 Db 278 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 297
 QY 573 GATACATAGCTATTGAAACACTAGCTATCGGTGACCAATACATCATCTCAAGAAATCTA 632
 Db 298 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 317
 QY 633 GCTCAAGCACAAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTATGACGTGAC 692
 Db 318 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 337
 QY 693 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 752

Db 338 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 357
 QY 753 TTTACTTACCGTGTAAAAATCGGAACAGCTTATAGGATCAATAAAAAATCTGGCTG 812
 Db 358 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 377
 QY 813 AATGAAGAAATAAACCACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 872
 Db 378 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 397
 QY 873 GAAACCGGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTACCATCAATAGCTT 932
 Db 398 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 417
 QY 933 GATGTCGATACCAACGAATTTGCTAAAAGTGACGAGCTCTTAACAGCTAGCGACGTAAC 992
 Db 418 AspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 437
 QY 993 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTTACTCTCAACAATCTC 1052
 Db 438 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 457
 QY 1053 GATGCTTTGCTATTATGACTATACCTTAACTGGAAGAGTAGAGGATATACGATGAC 1112
 Db 458 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 477
 QY 1173 CATTTAGCT 1181
 Db 498 HisLeuAla 500
 RESULT 10
 AAW94665
 ID AAW94665 standard; protein; 414 AA.
 XX
 AC AAW94665;
 XX
 DT 04-MAY-1999 (first entry)
 XX
 DE Streptococcus equisimilis mutant streptokinase K59E.
 XX
 KW Streptococcus equisimilis H46A; streptokinase; mutant; fibrinolytic;
 KW plasma clot; hydrolysis; haemolytic Streptococcus; plasminogen; plasmin;
 KW serine protease; fibrin; blood clot; thrombolytic;
 KW vascular thromboembolytic symptom; acute myocardial infarction;
 KW fibrinolysis; resistance.
 XX
 OS Streptococcus dysgalactiae subsp. equisimilis.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 14 /note= "encoded by ACC"
 PT
 XX US5876999-A.
 XX
 XX 02-MAR-1999.
 XX
 XX 06-DEC-1995; 95US-00568393.
 XX
 XX 06-DEC-1995; 95US-00568393.
 XX
 XX (NASC-) NAT SCI COUNCIL.
 XX
 XX Wu H;
 XX
 XX WPI; 1999-189643/16.
 DR N-PSDB; AAX16633.
 XX
 XX Mutant streptokinase polypeptide - useful as plasmin-resistant
 PT

DR WPI; 1991-208151/28.
 XX N-PSDB; AAQ12490.
 PT Fusion protein cleavage by blood clotting enzyme - for prodn. of
 PT fractions having greater antithrombotic activity for therapy and
 XX prophylaxis.
 PS Disclosure; Page 96; 115pp; English.
 XX
 CC The protein is a recombinant product of a gene fusion construct. The
 CC sequence of the synthetic hirudin HV-1 genes was designed based on the
 CC published amino acid sequence (Dodd J., et al FBS Letters 165 180
 CC (1984)). The sequence of streptokinase was obt'd. from PCR amplified
 CC chromosomal DNA from S. equisimilis ATCC 10009 or ATCC 9642. The primers
 CC used for the PCR were based on the pub- lished DNA sequence of S.
 CC equisimilis strain H46A (Malke, H., Roe, B., and Ferretti, J.J., Gene 34
 CC 357-362 [1985]). The two sequences were used to construct an expression
 CC vector in which the streptokinase gene is linked to the hirudin gene via
 CC a linking sequence encoding a cleavage site for factor Xa. The factor Xa
 CC is present at the site of the target thrombus so the active agents are
 CC released specifically at the place where clot formation is occurring. See
 CC also AAR12887-R12889, AAR12891-R12894 and AAR12885. (Updated on 25-MAR-
 CC 2003 to correct PA field.)
 XX
 SQ Sequence 483 AA;

Alignment Scores:
 Pred. No.: 4,46e-178 Length: 483
 Score: 1978.00 Matches: 401
 Percent Similarity: 82.28% Conservative: 17
 Best Local Similarity: 78.94% Mismatches: 52
 Query Match: 72.81% Indels: 38
 DB: 2 Gaps: 8

US-09-940-235-9 (1-1541) x AAR12522 (1-483)

QY 33 ATAGTGGTCTGTAATGGCTACTAGATCGTCTCTGTAATAACAGCCAAATGGTGTGTT 92
 DB 1 IleAlaGlyProGluTrpLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
 QY 93 AGCGTTGCTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAATC 152
 DB 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPheGluIle 40
 QY 153 GATCTAACATCAGACCTGCTCATGGAGGAAGACAGCAAGCTTAAGTCCAAATCA 212
 DB 41 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 60
 QY 213 AAACCATTTGCTAGTAGTGGCGCGATGTCCATATAACTTGAAGAACTGACTTACTA 272
 DB 61 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu 80
 QY 273 AAGGCTATTCAAGAACAAATTGATCGCTAACGTCACAGTAACGACGACTACTTTGAGGTC 332
 DB 81 LysAlaIleGlnGluLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100
 QY 333 ATTGATTTGCAAGCGATGCAACCAATTACTATGATCGAAACGCAAGCTCTACTTGTGAC 392
 DB 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
 QY 393 AAAGATGGTTCGGTAACTTCGCGACCACTGTCACAGAAATTTTGTAAAGCGGACAT 452
 DB 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
 QY 453 GTGCGGTTAGACCATATAAGAAAAACCAATACAAAAACCAAGCGAAATCTGTTGATGTG 512
 DB 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
 QY 513 GAATATACGTACAGTTTACTCCCTTAAACCTTGATGAGATTTTCAGACAGGTCTCAA 572
 DB 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuLys 180
 QY 573 GATACTAGCTATTGAAACACTACTACTATCGGTGACACCATCATCTCAAGNAATTACTA 632

DB 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 200
 QY 633 GCTCAAGCACAAAGCATTATAACAAAAACCAAGCAGCTATACGATTATGACGTCAC 692
 DB 201 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp 220
 QY 693 TCCTCAATCTGCTCACTCATGACATGACATGACATTTCCGTACGATTTTACCAATCGATCAAGAG 752
 DB 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
 QY 753 TTTACTTACCGTGTATAAATTCGGAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 812
 DB 241 PheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsnLysLysSerGlyLeu 260
 QY 813 AATGAGAATAAACAACACTGACCTGATCTCTGAGAAATATATAGTCTTTAAAAAGGG 872
 DB 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280
 QY 873 GAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCAACCATCAATACGTT 932
 DB 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
 QY 933 GATGTCGATACCAACGAATTCGTAAGAGTACGAGCTCTTAAACAGCTAGCGAAGCTAAC 992
 DB 301 AspValAsnThrAsnGluLeuLeuLysSerGluGlnLeuThrAlaSerGluArgAsn 320
 QY 993 TTAGACTTCAGAGATTTATACGATCTCTCGTGAATAGGCTAACTACTCTACAACAATCTC 1052
 DB 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
 QY 1053 GATGCTTTGCTGATATGAGTACTTAACTTAACTGGAAGCTAGAGGATTAATCAGATGAC 1112
 DB 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
 QY 1113 ACCAACGCTATACACCGTTTATATGGCAAGCGACCCGAGAGAGAGAAATGCTAGCTAC 1172
 DB 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 380
 QY 1173 CATTTAGCTGGTGGCCAGCGCCAAACAGATTGTACCCATGCTGAGAAAGTGTGTTGAT 1232
 DB 381 HisLeuAla-----TyrAspLysAspArgTyrThrGluGluGluArgGluValTyrSer 398
 QY 1233 CATGCTGCTGGGACTTCTATGTCGAGAAACGTCGAGAAAGCCCTACCAAGGCTGG 1292
 DB 399 TyrLeuArgTyrThrGlyThrProIleProAspAsnProAsnAspLysIleGluGlyArg 418
 QY 1293 ATGATG---GTAGATTGCTACT-----TGCCCTGGGAGAGAGGCGAGCGGA 1331
 DB 419 ValValTyrThrAspCysThrGluSerGlyGlnAsnLeuCysLeuCysGluGlySerAsn 438
 QY 1332 CGCATCATCTTGCACTTCTAGAAATAGATGCAACGATCAGGACACAAAGGACATCTCTATAGA 1391
 DB 439 -----ValCysGlyGlnGlyAsnLysCys-----TGCTGGGAGAGAGGAGGAGG 1447
 QY 1392 ATTGGAGACACCTGGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGCACA 1451
 DB 448 LeuGlySerAspGlyGluLysAsn-----GlnCysVal---Thr 459
 QY 1452 GGCACACGCGGAGGAGAGTGGAGGCGACACCTCTGTGACAGCCACACATCTGAGC 1511
 DB 460 GlyGluGly-----ThrProLysProGlnSerHis 469
 QY 1512 GGATCTGGCCCTTTCACCGATGTT 1535
 DB 470 AsnAspGlyAspPheGluGluIle 477
 RESULT 12
 AAW21728
 ID AAW21728 standard; protein; 813 AA.
 XX
 AC AAW21728;
 XX

Key Location/Qualifiers
FH 1.381
FT /label= Maltose binding protein.
FT /note= "acts as blocking group"
FT 382..1194
FT /label= Streptokinase
XX MO9641883-A1.
XX PN
XX XX
XX XX
XX PD 27-DEC-1996.
XX XX
XX 07-JUN-1996; 96WO-US009640.
XX XX
XX 09-JUN-1995; 95US-00488940.
XX XX
XX (HARD) HARVARD COLLEGE.
XX PA
XX XX
XX PI Reed GL;
XX XX
XX WPI; 1997-065469/06.
XX DR
XX XX
XX PT Modified forms of streptokinase resistant to enzymatic cleavage - useful
XX PT as thrombolytic agents in treating thrombosis and in medical equipment.
XX XX
XX PS Example 1; Page 11-12; 65pp; English.
XX CC This sequence represents a fusion protein between maltose-binding protein
XX CC and the plasminogen-binding fragment of streptokinase. This fusion
XX CC protein was used in the design of modified streptokinase has an in vitro
XX CC degradation rate at least 2 times slower than that of native
XX CC streptokinase. Compounds containing modified streptokinases are
XX CC specifically used as thrombolytic agents for dissolving blood clots in
XX CC vivo in a mammal, preferably at a dose of 2000 U/kg, optionally as a
XX CC bolus rather than by continuous infusion. (Updated on 17-OCT-2003 to
XX CC standardise OS field)
XX SQ Sequence 1194 AA;

Alignment Scores:
Pred. No.: 8.02e-177 Length: 1194
Score: 1966.50 Matches: 382
Percent Similarity: 99.74% Conservative: 0
Best Local Similarity: 99.74% Mismatches: 0
Query Match: 72.40% Indels: 1
DB: 2 Gaps: 1

US-09-940-235-9 (1-1541) x AAW21726 (1-1194)

QY 33 ATAGTGTGCTGTAATGGCTACTAGATCGTCTCTGTAAATAACAGCCAAATTTGGTTGT 92
DB 382 IleAlaGlyProGluTrpLeuAspArgProSerValAsnAsnSerGlnLeuValVal 401

QY 93 AGCGTTGCTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC 152
DB 402 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPheGluIle 421

QY 153 GATCTAACATCAGCCTGCTCATCGAGGAAGACAGCAAGCGTCTTAAGTCAAAATCA 212
DB 422 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 441

QY 213 AAACCATTTGCTAGTAGTGGCGGATGTCACATAAACTTGAGAAAGCTGACTACTA 272
DB 442 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 461

QY 273 AAGGCTATTCAAGAACAAATTGATCGCTAACGTCACAGTAACGACGACTACTTTGAGGTC 332
DB 462 LysAlaIleGlnGluLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 481

QY 333 ATTGATTTGCAAGCGATGCAACCAATTACTGATCGAAACGGCAAGTCTTACTTTGCTGAC 392
DB 482 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 501

QY 393 AAAGATGGTTCGGTAACCTTGCCGACCCCAACCTGTCCAGAAATTTTGTGTAAGCGGACAT 452

Db 502 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 521
QY 453 GTGCGGTTAGACCATATAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 512
Db 522 ValArgValArg--TyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 540
QY 513 GAATATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTCACACAGGCTCTCAAA 572
Db 541 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 560
QY 573 GATACCTAAGCTATTGAAACACATAGCTATCGGTGACACCATCATCTCAAGAAATACTA 632
Db 561 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 580
QY 633 GCTCAAGCACAAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTATGACAGTGAC 692
Db 581 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 600
QY 693 TCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 752
Db 601 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 620
QY 753 TTTACTTACCGTGTAAAAATCGGAAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 812
Db 621 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 640
QY 813 AATGAAGAATAAACAACTGACCTGATCTCGAGAAATATATTACGTCCTTAAAAAGGG 872
Db 641 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 660
QY 873 GAAACCGCTATGATCCCTTTGATCGCAGTCACCTGAAACTGTTTCAACCAATACGTT 932
Db 661 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 680
QY 933 GATGTCGATACCAACGAATTGCTAAAAAGTGACAGCTCTTAAACAGCTACGCAACGTAAC 992
Db 681 AspValAspThrAsnGlnLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 700
QY 993 TTAGACTTCAGAGATTTTATACGATCTCGTGATAAGCTAAACTCTCTCAACAATCTC 1052
Db 701 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 720
QY 1053 GATGCTTTGCTATTATGACTATACCTTAACTGGAAGAGTAGAGGATATACGATGAC 1112
Db 721 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 740
QY 1113 ACCAACCGTATCATACCCGTTTATATGCGCAACGACGCCCAAGAGAGAGATGCTAGCTAC 1172
Db 741 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 760
QY 1173 CATTTAGCT 1181
Db 761 HisLeuAla 763

RESULT 14
AEB93425
ID AEB93425 standard; protein; 415 AA.
XX
AC AEB93425;
XX
DT 20-OCT-2005 (first entry)
DE Streptococcus equisimilis streptokinase protein SEQ ID NO:1.
XX protein quantitation; streptokinase.
XX Streptococcus dysgalactiae subsp. equisimilis.
XX US2005176085-A1.
XX
PD 11-AUG-2005.
XX

15-NOV-2004; 2004US-00988943.
19-NOV-2003; 2003CU-00000269.
(BETA/) BETANCOURT N L H.
(DORT/) DORTA-DUQUE J F D C.
(PERE/) PEREZ V A B.
(VALD/) VALDES J G.
(LOPE/) LOPEZ L J G.
(PALO/) PALOMARES G R P.
(FEYR/) FEYR R P.
(GILF/) GIL F M A.
BETANCOURT NLH, Dorta-Duque JFD, Perez VAB, Valdes JG, Lopez LJG;
Palomares GRP, Feyr RP, Gil FMA;
WPI; 2005-596348/61.
Identifying and quantifying proteins in complex mixtures by selectively
isolating peptides not containing histidine nor arginine from each
protein, and determining the relative concentration of proteins in
different samples.
Example 3; SEQ ID NO 1; 91pp; English.
The invention relates to a method for identifying and quantifying one or
more proteins in complex mixtures by selectively isolating peptides not
containing histidine nor arginine (NHR peptides) from each protein, and
determining the relative concentration of one or more proteins in
different samples from the ratio between the areas of estimated
theoretical spectra for each NHR peptide labeled with different isotopes
in each sample. The method comprises: (a) enzymatically or chemically
hydrolyzing the sample or samples of proteins; (b) chemically modifying
alpha and epsilon amino groups (alpha- and epsilon-NH2) of every peptide
obtained in step (a); (c) isolating the NHR peptides by cation exchange
chromatography from the mixture of peptides obtained in step (b); (d)
identifying proteins by mass spectrometry analysis of the NHR peptides
obtained in step (c); (e) differential isotopic labeling of protein
samples previously to step (a) or during steps (a) or (b) and immediately
mixing at least a portion of the samples; and (f) relative quantifying of
more proteins in the mixtures of step (e) from the ratio between the
areas of estimated theoretical spectra of the pair of NHR peptides
identified in step (d), as well as from the ratio between the areas of
the estimated theoretical spectra of fragments from the NHR peptides,
generated in step (d). Also described is a kit for the identification and
quantification of proteins in complex mixtures, which comprises the
method mentioned above. The method and kit are useful for identifying or
quantifying proteins, such as those with vaccinal, therapeutic or
diagnostic uses, in complex mixtures. The present sequence represents a
streptokinase protein, which is used in an example from the present
invention.
SQ Sequence 415 AA;
Alignment Scores:
Pred. No.: 1.11e-176 Length: 415
Score: 1963.00 Matches: 379
Percent Similarity: 98.96% Conservative: 1
Best Local Similarity: 98.70% Mismatches: 4
Query Match: 72.28% Indels: 0
DB: 9 Gaps: 0
US-09-940-235-9 (1-1541) * AEB93425 (1-415)
QY 30 ATGATAGCTGCGTTCCTGAATGCGTCTTCTGTAAATTAACGCCAATTGGTT 89
DB 1 Mcttlltlaaglypndlturpneuapaygprosevalasnaabserglnleuval 20
QY 90 GTTAGGCTGCTGCTGCTGTTGAGGAGCAATCAAGACATTAAGTCTTAAATTTTGA 149
DB 21 valsevalalaglythrvalingulynrnsnglnasplleberleunyshpehgnlu 40
QY 150 ATCGATCTTAACATCAAGCCTGCTCAVTGAGGAGAAAGACAGAGCAGCTTAAGTCCAAA 209

ID AAW86143 standard; protein; 414 AA.
 XX AC AAW86143;
 XX
 DT 17-OCT-2003 (revised)
 DT 03-MAR-1999 (first entry)
 XX
 XX Streptokinase (SK) protein sequence.
 DE
 XX Non-immunogenic; epitope; T-cell; immunogenicity; immune system; SK;
 KW immunogl bulin; therapeutic; streptokinase.
 KW
 XX Streptococcus dysgalactiae subsp. equisimilis.
 OS
 XX WO9852976-A1.
 PN
 XX 26-NOV-1998.
 PD
 XX 21-MAY-1998; 98WO-GB001473.
 PF
 XX 21-MAY-1997; 97GB-00010480.
 PR 31-JUL-1997; 97GB-00016197.
 PR 28-NOV-1997; 97GB-00025270.
 PR 02-DEC-1997; 97US-0067235P.
 PR 14-APR-1998; 98GB-00007751.
 PR
 XX (BIOV-) BIOVATION LTD.
 PA
 XX Carr FU;
 PI
 XX WPI; 1999-045301/04.
 XX
 XX Reducing immunogenicity of proteins - by modifying the amino acid
 PT sequence of the protein to eliminate potential epitopes for T-cells of a
 PT given species.
 PT
 XX Example 6; Fig 28; 77pp; English.
 PS
 XX The invention relates to a method for the production of non-immunogenic
 CC proteins. The method comprises determining at least part of the amino
 CC acid sequence of the protein; (b) identifying in the amino acid sequence
 CC one or more potential epitopes for T-cells (T-cell epitopes) of the given
 CC species; and (c) modifying the amino acid sequence to eliminate at least
 CC one of the T-cell epitopes identified in step (b) thereby to eliminate or
 CC reduce the immunogenicity of the protein when exposed to the immune
 CC system of the given species. A method of analysing a pre-existing protein
 CC to predict the basis for immunogenic responses is also provided. The
 CC methods can be used particularly for reducing the immunogenicity of
 CC immunoglobulins or therapeutic proteins, e.g. Streptokinase (SK). The
 CC products can be used for diagnosis and therapy. The present sequence
 CC represents the amino acid sequence of the SK protein. (Updated on 17-OCT-
 CC 2003 to standardise OS field)
 XX
 SQ Sequence 414 .AA;

Alignment Scores:
 Pred. No.: 3,29e-176 Length: 414
 Score: 1958.00 Matches: 378
 Percent Similarity: 98.96% Conservative: 1
 Best Local Similarity: 98.69% Mismatches: 4
 Query Match: 72.09% Indels: 0
 DB: 2 Gaps: 0

US-09-940-235-9 (1-1541) x AAW86143 (1-414)

QY 33 ATAGCTGGTCCTGAATGGCTACTAGATCGCTCTCTGTAATAACAGCCAAATTTGGTTGTT 92
 Db 1 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal 20

QY 93 AGCGTGTGCTGTCTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAATC 152
 Db 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 40

QY 153 GATCTAACATCAGCACTGCTCATGGAGGAAAGACAGACAGCAAGCTTAAAGTCAAAATCA 212
 Db 41 AspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 60

QY 213 AAACCAATTTGCTACTGATAGTGGCGGATGTCATAAACTTGAGAAAGCTGACTTACTA 272
 Db 61 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu 80

QY 273 AAGCTATTCAAGACAAATTTGATCGGTAAAGTCCACAGTAACGACGACTTCTTGAGTTC 332
 Db 81 LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100

QY 333 ATTGATTTTTCGAAGCATGCAACATTACTGATGCGAAACGCGCAAGGCTCTACTTTGCTGAC 392
 Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120

QY 393 AAAGATGGTTCGGTAACCTTTGCCGACCCAACTGTCTCAAGAATTTTTCGTAAGCGGACAT 452
 Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140

QY 453 GTCCGGGTAGACCATATAAGAAAACCAATACAAAACCAAGGAAATCTGTGATGTG 512
 Db 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160

QY 513 GAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAA 572
 Db 161 GlutyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180

QY 573 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAATTTACTA 632
 Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 200

QY 633 GCTCAAGCAACAAAGCATTTTAAACAAAACCAACCCAGGCTATAGATTTATGAACGTGAC 692
 Db 201 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp 220

QY 693 TCCTCAATCGTCACTCATGACAAATGATTTTCGTCAGATTTTACCAGTGGATCAAGAG 752
 Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240

QY 753 TTTTACTTACCGTGTAAAAATCGGAAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 812
 Db 241 PheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsnLysLysSerGlyLeu 260

QY 813 AATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTAACGTCCTTAAAAAGGG 872
 Db 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280

QY 873 GAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTCACCATCAATACGTT 932
 Db 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300

QY 933 GATGTCGATACCAACGAATTTGCTAAAAGTGAGCAGCTCTTAAACGCTAGCAGCACTAAC 992
 Db 301 AspValAsnThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320

QY 993 TTAGACTTCAGACATTTATACGATCTCTGCGTGAAGGCTAAACTACTCTCAACAATCTC 1052
 Db 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340

QY 1053 GATGCTTTGGTATTATGGACTATACCTTAACTGGAAAAGTAGAGGATATACGATGAC 1112
 Db 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360

QY 1113 ACCAACCGTATACACCGTTTATATGGGCAAGCGACCCGAGGAGAGAGATGCTAGCTAC 1172
 Db 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 380

QY 1173 CATTTAGCT 1181
 Db 381 HisLeuAla 383

Search completed: January 28, 2006, 02:08:00
Job time : 126.195 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 28, 2006, 01:49:55 ; Search time 21.9832 Seconds
(without alignments)
11590.966 Million cell updates/sec

Title: US-09-940-235-9

Perfect score: 2716

Sequence: 1 ttgttacttaagaag.....cctcacgagtgctgtag 1541

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 1144120

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09940235/raunat_27012006_144219_27602/app_query.fasta_1.7708
-DB=issued_Patents_AA -QFAST=fastan -SUFFIX=rai -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsnm62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09940235@cgn2_1.159 @raunat_27012006_144219_27602 -NCPU=6 -ICPU=3
-NO_MMAP -LARGESQUERY -NEG_SCORES=30 -THREADS=1 -XGAPOP=10 -LONGIOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: Issued_Patents_AA:*
2: /cgn2_6/ptodata/1/1aa/5_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/H_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/RE_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfillset1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1984	73.0	413	2	US-10-360-101-264 Sequence 264, App
2	1984	73.0	414	6	US-09-211-542A-6 Sequence 6, Appl1
3	1984	73.0	414	6	5240845-1 Patent No. 5240845
4	1984	73.0	440	1	US-08-560-098A-52 Sequence 52, Appl1
5	1984	73.0	795	2	US-09-211-542A-2 Sequence 2, Appl1
6	1978	72.8	483	1	US-07-854-596B-47 Sequence 47, Appl1
7	1966.5	72.4	813	1	US-08-488-940-3 Sequence 3, Appl1
8	1966.5	72.4	1194	1	US-08-488-940-1 Sequence 1, Appl1
9	1963	72.3	435	1	US-07-854-596B-26 Sequence 26, Appl1
10	1958	72.1	440	1	US-07-854-596B-19 Sequence 19, Appl1
11	1958	72.1	440	1	US-07-854-596B-15 Sequence 15, Appl1
12	1958	72.1	483	1	US-07-854-596B-43 Sequence 43, Appl1

13	1958	72.1	499	1	US-07-854-596B-28 Sequence 28, Appl1
14	1958	72.1	859	1	US-07-854-596B-35 Sequence 35, Appl1
15	1937.5	71.3	1194	1	US-08-488-940-18 Sequence 18, Appl1
16	1936.5	71.3	1194	1	US-08-488-940-17 Sequence 17, Appl1
17	1928	71.0	413	1	US-08-759-559-12 Sequence 12, Appl1
18	1928	71.0	413	1	US-09-294-457-12 Sequence 12, Appl1
19	1928	71.0	413	2	US-09-919-703-12 Sequence 12, Appl1
20	1902	70.0	384	2	US-09-374-038-4 Sequence 4, Appl1
21	1902	70.0	384	2	US-09-658-179-4 Sequence 4, Appl1
22	1900	70.0	800	1	US-08-488-940-4 Sequence 4, Appl1
23	1900	70.0	1181	1	US-08-488-940-2 Sequence 2, Appl1
24	1898	69.9	372	2	US-09-374-038-3 Sequence 3, Appl1
25	1898	69.9	372	2	US-09-658-179-3 Sequence 3, Appl1
26	1890.5	69.6	747	1	US-07-854-596B-40 Sequence 40, Appl1
27	1887	69.5	401	2	US-09-374-038-1 Sequence 1, Appl1
28	1887	69.5	401	2	US-09-658-179-1 Sequence 1, Appl1
29	1887	69.5	413	2	US-09-374-038-2 Sequence 2, Appl1
30	1887	69.5	413	2	US-09-658-179-2 Sequence 2, Appl1
31	1875	69.0	369	1	US-07-854-596B-31 Sequence 31, Appl1
32	1695.5	62.4	736	2	US-09-211-542A-4 Sequence 4, Appl1
33	1686	62.1	356	2	US-09-211-542A-12 Sequence 12, Appl1
34	785	28.9	150	2	US-09-211-542A-14 Sequence 14, Appl1
35	632	23.3	2231	1	US-08-153-799-16 Sequence 16, Appl1
36	632	23.3	2324	1	US-08-283-857-1 Sequence 1, Appl1
37	632	23.3	2324	4	PCT-US95-09819-1 Sequence 1, Appl1
38	632	23.3	2355	2	US-10-360-101-235 Sequence 235, App
39	632	23.3	2386	1	US-09-016-366A-12 Sequence 12, Appl1
40	632	23.3	2386	2	US-09-961-403-1 Sequence 1, Appl1
41	632	23.3	2446	1	US-08-551-356-2 Sequence 2, Appl1
42	632	23.3	2446	4	PCT-US93-12687-2 Sequence 2, Appl1
43	627	23.1	2327	6	5455158-1 Patent No. 5455158
44	283	10.4	286	2	US-09-078-091-4 Sequence 4, Appl1
45	275	10.1	286	2	US-09-078-091-2 Sequence 2, Appl1

ALIGNMENTS

RESULT 1
US-10-360-101-264
; Sequence 264, Application US/10360101
; Patent No. 6861236
; GENERAL INFORMATION:
; APPLICANT: Moll, Gert N
; TITLE OF INVENTION: Export and modification of (poly)peptide in the antibiotic way
; FILE REFERENCE: 2183-5673
; CURRENT APPLICATION NUMBER: US/10/360,101
; CURRENT FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: BP 02077060.8
; PRIOR FILING DATE: 2002-05-24
; NUMBER OF SEQ ID NOS: 309
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 264
; LENGTH: 413
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: sequence of streptokinase
US-10-360-101-264

Alignment Scores:

Pred. No.:	2.88e-192	Length:	413
Score:	1984.00	Matches:	383
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	73.05%	Indels:	0
DB:	2	Gaps:	0

US-09-940-235-9 (1-1541) x US-10-360-101-264 (1-413)

Qy	33	ATAGCGTCCTGTAAGGCTACTAGATCGCTTCTGTAATAACAGCAATGCTGTT	92
Db	1	11a1a1g1ygcg1utp1a1e1u1a1p1a1r1p1r1a1s1a1n1s1e1r1g1n1e1u1a1v1a1	20

```
QY 93 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAATC 152
Db |||||||
21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPheGluIle 40
QY 153 GATCTAACATCAGCACTGCTCATGGAGGAAAGACAGACCAAGGCTTAAAGTCCAAATCA 212
Db |||||||
41 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 60
QY 213 AAACCATTTGCTACTGATAGTGGCGGATGTACATAAACTTGAGAAAGCTGACTTACTA 272
Db |||||||
61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 80
QY 273 AAGGCTATTCAAGAACAAATTGATCGCTAACGCTCCACAGTAACGACGACTACTTTGAGTTC 332
Db |||||||
81 LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspAspIlePheGluVal 100
QY 333 ATTGATTTTGAAGCGATGCAACCACTTACTGATCGAAACGGCAAGGCTACTTTGCTGAC 392
Db |||||||
101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValIlePheAlaAsp 120
QY 393 AAAGATGTTGCTGTAACCTTGGCGGACCACTGCTCCAAAGAATTTTGTCTAAGCGGACAT 452
Db |||||||
121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
QY 453 GTGCGCGTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 512
Db |||||||
141 ValArgValArgProTyrLysGlnLysProIleGlnAsnGlnAlaLysSerValAspVal 160
QY 513 GAATATCTGTACAGTCTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTCAAA 572
Db |||||||
161 GluTyrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuLys 180
QY 573 GATCTAAGCTATTGAAACACTAGCTATCGGTGACCATCATCATCTCAAGAATTACTA 632
Db |||||||
181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 200
QY 633 GCTCAAGCACAAAGCATTTTAAACAAACCCAGGCTATACGATTTTATGACGTCGAC 692
Db |||||||
201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 220
QY 693 TCCTCAATCGTCACATCATGACATGACATTTTCGCTACGATTTTACCAATGGATCAAGAG 752
Db |||||||
221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
QY 753 TTTACTTACCGTGTAAAATCGGGAACAAGCTTATPAGGATCAATAAAAAATCTGCTCTG 812
Db |||||||
241 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 260
QY 813 AATGAGAAATAAACAACACTGACCTGATCTCTGAGAAATATATAGCTCTTAAAAAGGG 872
Db |||||||
261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280
QY 873 GAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCACCATCAAAATCGTT 932
Db |||||||
281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
QY 933 GATGTCGATACCAACGAATTGCTTAAAGTGTAGCAGCTCTTAAACAGCTAGCGAACGTAAC 992
Db |||||||
301 AspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320
QY 993 TTAGACTTCAGAGATTATACATCTCTGATAGGCTAAAGCTAACTACTCTACAAATCTC 1052
Db |||||||
321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
QY 1053 GATGCTTTTGGTATTATGACTATATCTTAACTTGAAGAGTAGAGGATAATACGATGAC 1112
Db |||||||
341 AspAlaPheGlyIleWeAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
QY 1113 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAGGAGAGAAATGCTAGCTAC 1172
Db |||||||
361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 380
```

```
QY 1173 CATTAGCT 1181
Db |||||||
381 HisLeuAla 383
RESULT 2
US-09-211-542A-6
; Sequence 6, Application US/09211542A
; Patent No. 6210667
; GENERAL INFORMATION:
; APPLICANT: Reed, Guy L.
; TITLE OF INVENTION: BACTERIAL FIBRIN-DEPENDENT PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROMBERG & SUNSTEIN, LLP
; STREET: 125 Summer Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/211,542A
; FILING DATE: 15-December-1998
; CLASSIFICATION: 1653
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/069,497
; FILING DATE: 15-December-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Attorney, Strimpel, Harriet M.
; REGISTRATION NUMBER: 37,008
; REFERENCE/DOCKET NUMBER: 1874/111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 443-9292
; TELEFAX: (617) 443-0004
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 414 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-211-542A-6
Alignment Scores:
Pred. No.: 2,89e-192 Length: 414
Score: 1984.00 Matches: 383
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 73.05% Indels: 0
DB: 2 Gaps: 0
US-09-940-235-9 (1-1541) x US-09-211-542A-6 (1-414)
QY 33 ATAGCTGGTCTCAATGGCTACTAGATCGCTCTTCTGTAATAATACAGCCAAATGGTGT 92
Db |||||||
1 IleAlaGlyProGluIleThrLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
QY 93 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAATC 152
Db |||||||
21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPheGluIle 40
QY 153 GATCTAACATCAGCACTGCTCATGGAGGAAAGACAGACCAAGGCTTAAAGTCCAAATCA 212
Db |||||||
41 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 60
QY 213 AAACCATTTGCTACTGATAGTGGCGGATGTACATAAACTTGAGAAAGCTGACTTACTA 272
Db |||||||
61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 80
QY 273 AAGGCTATTCAAGAACAAATTGATCGCTAACGCTCCACAGTAACGACGACTACTTTGAGTTC 332
```

```

Db      ||| 81 LysAlaIleGlnGlnLeuIleAlaAsnValHisSerAsnAspTyrPheGluVal 100
QY      ||| 333 ATTGATTTGGCAAGCGAATGCAACATTACGATCGAAACGGCAAGCTCTTGGTCGAC 392
Db      ||| 101 IleAspPheAlaSerAlaThrIleThrAspArgAsnGlyValTyrPheAlaAsp 120
QY      ||| 393 AAGATGATTCGGTAACTTGGCCGACCACTGTCGCAAGAAATTTGGTAAAGCGGACAT 452
Db      ||| 121 LysAspGlySerAlaThrLeuProthGlnProValGlnGlnPheLeuLeuSerGlyHis 140
QY      ||| 453 GTGGCGGTAAAGACCAATTAAGAAAAACAATCAAAAACAAGGAAATCTGTTGATGTG 512
Db      ||| 141 ValArgValArgProTyrLysGlnLysProIleGlnAsnGlnAlaLysSerValAspVal 160
QY      ||| 513 GAATATACGTGTACGTTTACCTCCCTTAAACCTGATGACGATTTCAACAGGCTGTA 572
Db      ||| 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
QY      ||| 573 GATATCAAGCTATTGAAACACTAGCTATCGGTGACACATCACTCATCTCAAGAACTACTA 632
Db      ||| 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeu 200
QY      ||| 633 GCTCAAGCAACAAGCAATTTTAAACAACCAACCCAGGCTATACGATTTTGAACGTGAC 692
Db      ||| 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGlnLysArg 220
QY      ||| 693 TCCCTCAATGCTCACTGATGACATGACATTTTCCGTCAGATTTTACCAATGATCAAGAG 752
Db      ||| 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
QY      ||| 753 TTACTTACCGTGTAAATGCGGAACAAGCTTATAGATCAATAAATCGTGTG 812
Db      ||| 241 PheThrTyrArgValLysAsnArgGlnGlnAlaTyrArgIleAsnLysLysSerGlyLeu 260
QY      ||| 813 AATGAAGAATAAACAACAACCTGACCTGATCTGAGAAATATTACGTCCTTAAACCAAGG 872
Db      ||| 261 AsnGlnGlnIleAsnAsnThrAspLeuIleSerGlnLysTyrTyrValLeuLysGly 280
QY      ||| 873 GAAAGACCGGTATGATCCCTTTGATGCGAGTCACTTGAACCTGTTCAACCAATACGTT 932
Db      ||| 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
QY      ||| 933 GATTCGATACCAACCAATTCGTTAAAGAGACAGCTCTTAACAGTACGGAACGTAAC 992
Db      ||| 301 AspValAspThrAsnGlnLeuLeuLysSerGlnGlnLeuThrAlaSerGlnLysArgAsn 320
QY      ||| 993 TTAGACTTCAGAGATTATACGATCTCGGTATAGGCTTAACTACTCTCAACAATCTC 1052
Db      ||| 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
QY      ||| 1053 GATGCTTTGGTATTATGACATATACCTTAACCTGAAAGATAGAGATTAATCAAGATGAC 1112
Db      ||| 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGlnAspAsnHisAspAsp 360
QY      ||| 1113 ACCAACCGTATCATTAACCGTTTATATGGGCAAGCGACCCGAGAAGAGAAATGCTAGTAC 1172
Db      ||| 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGlnGlnAlaSerTyr 380
QY      ||| 1173 CATTTAGCT 1181
Db      ||| 381 HisLeuAla 383

```

```

; FILING DATE: 06-JUL-1990
; SEQ ID NO: 1
; LENGTH: 414
5240845-1

Alignment Scores:
Pred. No.: 2,896-192 Length: 414
Score: 1984.00 Matches: 383
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 73.05% Indels: 0
DB: 6 Gaps: 0

US-09-940-235-9 (1-1541) x 5240845-1 (1-414)

QY      ||| 33 ATAGCTGCTGCTGAATGGCTACTAGATGTGCTCTGTAAATTAACGCCAATGGTGT 92
Db      ||| 1 IleAlaGlyProGlnTyrPheLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
QY      ||| 93 AGCGTGTGCTGCTGTTGAGGGGCAAGCAATTAAGCTTAAATTTTGAATC 152
Db      ||| 21 SerValAlaGlyThrValGlnGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 40
QY      ||| 153 GATTAACATCAAGCACTGCTCATGAGGAAAGACAGCAAGGCTTAAGTCAAAATCA 212
Db      ||| 41 AspLeuThrSerArgProAlaHisGlyGlyLysThrGlnGlnGlyLeuSerProLysSer 60
QY      ||| 213 AAACCAATTTGCTACTGATGATGGCGGATGTCACATTAACCTTGAAGAACTGACTACTA 272
Db      ||| 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlnLysValAspLeuLeu 80
QY      ||| 273 AAGGCTATTCAGAAACAATGATGGCTTACGTCACAGTAACGAGTACTTGAAGGTC 332
Db      ||| 81 LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100
QY      ||| 333 ATTGATTTGGCAAGCGAATGCAACATTACGATCGAAACGGCAAGCTCTTGGTCGAC 392
Db      ||| 101 IleAspPheAlaSerAlaThrIleThrAspArgAsnGlyValTyrPheAlaAsp 120
QY      ||| 393 AAGATGATTCGGTAACTTGGCCGACCACTGTCGCAAGAAATTTGGTAAAGCGGACAT 452
Db      ||| 121 LysAspGlySerAlaThrLeuProthGlnProValGlnGlnPheLeuLeuSerGlyHis 140
QY      ||| 453 GTGGCGGTAAAGACCAATTAAGAAAAACAATCAAAAACAAGGAAATCTGTTGATGTG 512
Db      ||| 141 ValArgValArgProTyrLysGlnLysProIleGlnAsnGlnAlaLysSerValAspVal 160
QY      ||| 513 GAATATACGTGTACGTTTACCTCCCTTAAACCTGATGACGATTTCAACAGGCTGTA 572
Db      ||| 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
QY      ||| 573 GATATCAAGCTATTGAAACACTAGCTATCGGTGACACATCACTCATCTCAAGAACTACTA 632
Db      ||| 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeu 200
QY      ||| 633 GCTCAAGCAACAAGCAATTTTAAACAACCAACCCAGGCTATACGATTTTGAACGTGAC 692
Db      ||| 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGlnLysArg 220
QY      ||| 693 TCCCTCAATGCTCACTGATGACATGACATTTTCCGTCAGATTTTACCAATGATCAAGAG 752
Db      ||| 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
QY      ||| 753 TTACTTACCGTGTAAATGCGGAACAAGCTTATAGATCAATTAATTAATGCTGTG 812
Db      ||| 241 PheThrTyrArgValLysAsnArgGlnGlnAlaTyrArgIleAsnLysLysSerGlyLeu 260
QY      ||| 813 AATGAAGAATAAACAACAACCTGACCTGATCTGAGAAATATTACGTCCTTAAACCAAGG 872
Db      ||| 261 AsnGlnGlnIleAsnAsnThrAspLeuIleSerGlnLysTyrTyrValLeuLysGly 280
QY      ||| 873 GAAAGACCGGTATGATCCCTTTGATGCGAGTCACTTGAACCTGTTCAACCAATACGTT 932

```

```

RESULT 3
5240845-1
; PATENT NO. 5240845
; APPLICANT: FUJII, SETSURO; TAKADA, KAORUKO.; KATANO, TAMIKI;
; MAJIMA, EIJI; OGINO, KOICHI; ONO, KENJI; SAKATA, YASUYO; UENOYAMA,
; TSUTOMU
; TITLE OF INVENTION: MUTATED STREPTOKINASE PROTEINS
; NUMBER OF SEQUENCES: 65
; CURRENT APPLICATION DATA.
; APPLICATION NUMBER: US/07/549,049

```

Db 281 GluysProTyraSProheaspargSerHisleuylsleuPheThrIleuTyTyrVal 300
 QY 933 GATGTCGATACCAAGCAATTTGGCTAAAGAGACAGCTCTTAACAGCTAGCGCAACCTAAC 992
 Db 301 AspValAspThrAsnGluIleuLeuIleuSerGluGlnIleuLeuThrAlaSerGluAsn 320
 QY 993 TTAGACTTCAGAGATTTATACGATCTCTCGTATAGGCTAAACTACTTACCAAACTCTC 1052
 Db 321 LeuAspPheArgAspLeuTyraSProArgAspLysAlaLysleuLeuTyraSAsnIleu 340
 QY 1053 GATGCTTTGGTATTATGCACTATACCTTAACCTGAGAAAGTAGAGATTAATCAAGATGAC 1112
 Db 341 AspAlaPheGlyIleMetAspTyThrIleuThrGlyLysValGluAspAsnHisAspAsp 360
 QY 1113 ACCAACCGATACATACCGTTTATATAGGCAAGCGAACCGAGAGAGAAATGCTAGCTAC 1172
 Db 361 ThrAsnArgIleIleThrValIleuThrGlyLysArgProGluGluAsnAlaSerTyTyr 380
 QY 1173 CATTAGCT 1181
 Db 381 HisIleuAla 383
 RESULT 4
 US-08-560-098A-52
 : Sequence 52, Application US/08560098A
 : Patent No. 5976841
 : GENERAL INFORMATION:
 : APPLICANT: WNEBNT, Stephan
 : APPLICANT: HEINZEL, WIELAND, Regina
 : APPLICANT: STEFFENS, Gerd Josef
 : TITLE OF INVENTION: Proteins having fibrinolytic and
 : NUMBER OF INVENTION: Coagulation-inhibiting Properties
 : CORRESPONDENCE ADDRESSES: 60
 : ADDRESSER: Evenston, McKee, Edwards & Lenahan
 : STREET: 1200 G Street, N.W., Suite 700
 : CITY: Washington
 : STATE: DC
 : COUNTRY: USA
 : ZIP: 20005
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/560,098A
 : FILING DATE: 17-NOV-1995
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: P 44 40 892.7
 : FILING DATE: 17-NOV-1994
 : ATTORNEY/AGENT INFORMATION:
 : NAME: EVANS, Joseph D.
 : REGISTRATION NUMBER: 26,269
 : REFERENCE/DOCKET NUMBER: 146/42448
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (202) 628-8800
 : TELEFAX: (202) 628-8844
 : INFORMATION FOR SEQ ID NO: 52:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 440 amino acids
 : TYPE: amino acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear
 : MOLECULE TYPE: protein
 : US-08-560-098A-52
 Alignment Scores:
 Pred. No.: 2,98e-192 Length: 440
 Score: 1984.00 Matches: 383
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 73.05% Indels: 0

DB: 1 Gaps: 0
 US-09-940-235-9 (1-1541) x US-08-560-098A-52 (1-440)
 QY 33 ATAGCTGGTCTCTGAATGGCTACTAGATCGTCTCTTCTGTAAATTAACACCAATGGTGT 92
 Db 27 IleAlaGlyProGluThrIleuLeuAspArgProSerValAsnSerGlnIleuVal 46
 QY 93 AGCGTGGCTGGTCTGTGGTGGGAGCAATCAAGACACTTGTCTAAATTTTGAATC 152
 Db 47 SerValAlaGlyThrValGluGlyThrAsnGlnAspLleSerLeuLysPheGlnIle 66
 QY 153 GATCTAACATCAGACCGCTGCTCATGAGAGAAAGACAGAGCAAGCTTAAGTCAAAATCA 212
 Db 67 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnIleuSerProLysSer 86
 QY 213 AAACCATTTGCTACGTAGTATGCGCGCATGATCATTAACCTTGAGAAAGCTGACTTACTA 272
 Db 87 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeu 106
 QY 273 AAGGCTATTCAGAACAAATTTGATCGCTAACGTCACAGTAAAGCACTACTTTGAGATC 332
 Db 107 LysAlaIleGlnIleuGlnIleuIleAlaAsnValHisSerAsnAspTyTyrPheGluVal 126
 QY 333 ATGATTTTTCAGACGATGCAACCATTAAGTATGATGCAAGCAAGCTACTTTGCTGAC 392
 Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyTyrPheAlaAsp 146
 QY 393 AAAGATGTTTCGGTAACTTTGCCGACCCCAACCTGTCCAGAAATTTTGTAAAGCGACAT 452
 Db 147 LysAspGlySerValThrIleuProThrGlnProValGlnIleuPheLeuSerGlyHis 166
 QY 453 GTCCGGCTTACACCTATTAAGAAACCAATCAAAACCAAGCAAGAAATCTGTGAATGTC 512
 Db 167 ValArgValArgProTyTyrGluLysProIleGlnAsnGlnAlaLysSerValAspVal 186
 QY 513 GAATATACCTGTACAGTTTACCTCCCTTAAACCTGATGACAGTTTCAGACAGGCTCAAA 572
 Db 187 GluTyThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 206
 QY 573 GATATACGATATTGAAAACATACATGATCGGTGACACATCAATCTCAAGATTACTA 632
 Db 207 AspThrLysleuLeuLysThrIleuAlaIleGlyAspThrIleThrSerGlnIleuLeu 226
 QY 633 GCTCAAGCACAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTGAACGTGAC 692
 Db 227 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyThrIleTyGluArgAsp 246
 QY 693 TCCTCAATGCTCATCATGACATGACATTTTCCGTACGATTTTAAACCAATGATCAAGAG 752
 Db 247 SerSerIleValThrHisAspAsnAspLlePheArgThrIleuProMetAspGlnGlu 266
 QY 753 TTACTTACCGTGTAAATTCGGGAAACAAGCTTATAGATCAATAAATATCGTCTG 812
 Db 267 PheThrTyArgValLysAsnArgGluGlnAlaIleTyArgIleAsnLysLysSerGlyLeu 286
 QY 813 AATGAAGAAATTAACAACATGACCTGATGCTGAGAAATTTTACGCTTTAAAAAGG 872
 Db 287 AsnGluGluIleAsnThrAspLeuIleSerGluLysTyTyrValIleuLysGly 306
 QY 873 GAAAGCCGATGATCCCTTTGATCGAGACATGTAACCTGTCACATCAATCAATAGTGT 932
 Db 307 GluLysProTyraSProPheAspArgSerHisleuLysLeuPheThrIleLysTyTyrVal 326
 QY 933 GATGTCGATACCAAGCAATTTGCTAAAGTAGACAGCTCTTAAACAGCTAGCGAAGCTAAC 992
 Db 327 AspValAspThrAsnGluIleuLeuLysSerGluGlnIleuLeuThrAlaSerGluAsn 346
 QY 993 TTAGACTTCAGAGATTTATACGATCTCTCGTATAGGCTAAACTACTTACCAAACTCTC 1052
 Db 347 LeuAspPheArgAspLeuTyraSProArgAspLysAlaLysleuLeuTyraSAsnIleu 366
 QY 1053 GATGCTTTGGTATTATGCACTATACCTTAACCTGAGAAAGTAGAGATTAATCAAGATGAC 1112

|||||
Db 367 AsphalaPheGlyIleMetAspYrThrLeuThrgIySValGIuAspAsnHisAspAsp 386
Qy 1113 ACCAACCGTATCATTAACCGTTTATATAGGCAAGCAGCCGGAAGAGGAATGCTAGCTAC 1172
Db 387 ThrAsnArgIleIleThrValIyrMeGlyLysArgProGIuGIuAsnAlaSerIyr 406
Qy 1173 CATTTAGCT 1181
Db 407 HisLeuAla 409
RESULT 5
US-09-211-542A-2
; Sequence 2, Application US/09211542A
; Patent No. 6210667
; GENERAL INFORMATION:
; APPLICANT: Reed, Guy L.
; TITLE OF INVENTION: BACTERIAL FIBRIN-DEPENDENT PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROBERG & SUNSTEIN, LLP
; STREET: 125 Summer Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/211,542A
; FILING DATE: 15-December-1998
; CLASSIFICATION: 1653
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/069,497
; FILING DATE: 15-December-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Attorney, Scrimpel, Harriet M.
; REGISTRATION NUMBER: 37,008
; REFERENCE/DOCKET NUMBER: 1874/111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)443-9292
; TELEFAX: (617)443-0004
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 795 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-211-542A-2
Alignment Scores:
Pred. No.: 3,99e-192 Length: 795
Score: 1984.00 Matches: 383
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 73.05% Indels: 0
DB: 2 Gaps: 0
US-09-940-235-9 (1-1541) x US-09-211-542A-2 (1-795)
Qy 33 ATAGCTGTGCTCGAATGCTACTAGATCGTCTTCTGTAAATAACAGCCAAATGTTGTT 92
Db 362 IleAlaGlyProGIuIyrIleuLeuAspArgProSerValAsnHisSerGIuLeuVal 401
Qy 93 AGCGTTGCTGTACTGTTTGAAGGAGCAATCAAGACTTAAATTTTGAATC 152
Db 402 SerValAlaGlyThrValGIuGIuIyrThrAsnGIuAspIleSerLeuIyrSphenGluIle 421
Qy 153 GATCTAACTACGACAGCTGCTCATGAGGAAGAAGAGCAAGCTTAATCCAAATCA 212
|||||

Db 422 AspleuThrSerArgProAlaHisGIyGIyIyrThrgIuGIuGIuLeuSerProIyrSer 441
Qy 213 AACCATTTGGCTACTAGATGAGCGGATGTCATTAACCTTGAAGAAAGCTGACTACTA 272
Db 442 LysProPheAlaThrAspSerGIyAlaMetSerHisLysLeuGIuLysAlaAspLeuLeu 461
Qy 273 AAGCTATTCAGAAACAATTGATCGCTAACGTCACAGTAAACGACGACTACTTTGAGGTC 332
Db 462 LysAlaIleGIuGIuGIuLeuIleAlaAsnValHisSerAsnAspAspIyrPheGIuVal 481
Qy 333 ATTGATTTTGCAGAGGATGCAACCATTTACTGATCGAAACGGCAGAGTCTACTTGTGAC 392
Db 482 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGIyLysValIyrPheAlaAsp 501
Qy 393 AAAGATGCTGGTAACTTGCACCGCAACCTGTCGAAAGATTCTTGTAAACCGGACAT 452
Db 502 LysAspGIySerValThrLeuProThrGIuProValGIuGIuPheLeuLeuSerGIyHis 521
Qy 453 GTGGCGCTTAGACATATATAAGAAACCAATATACAAACCAAGCAAAATCTGTTGATGTG 512
Db 522 ValArgValArgProIyrLysGIuLysProIleGIuAsnGIuAlaLysSerValAspVal 541
Qy 513 GAATATACTGTACAGTTTACTCTCTTAACCTGTATGACGATTTGACACAGTCTCAAA 572
Db 542 GIuIyrThrValGIuPheThrProLeuAsnProAspAspAspPheArgProGIyLeuLys 561
Qy 573 GATACTAAGCTATTTGAAACACTAGCTATCGGTGACACCATTCACACTTCAGAAATTACTA 632
Db 562 AspThrLysLeuLeuLysThrLeuAlaIleGIyAspThrIleThSerGIuLeuLeu 581
Qy 633 GCTCAAGACAAACATTTTAAACAAAACCAACCCAGGCTTATACATTTTATGACAGTAC 692
Db 582 AlaGIuAlaGlnSerIleLeuAsnLysAsnHisSProGIyIyrThIleIyrGIuArgAsp 601
Qy 693 TCCTCAATCGTCACTGACATGACATGACATTTTCGTAGCAATTTTACCAATGATCAAGAG 752
Db 602 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGIuGIu 621
Qy 753 TTTTACTTACCGTGTAAAAATCGGGAACAAGCTTATAGATCAATAAAAATCGGTGTG 812
Db 622 PheThrIyrArgValLysAsnArgGIuGIuAlaIyrArgIleAsnLysLysSerGIyLeu 641
Qy 813 AATGAGAAATTAACAACAACGACCTGATCTGTAGAAATTTACGCTTAAAAAGAG 872
Db 642 AsnGIuGIuIleAsnAsnThrAspLeuIleSerGIuLysIyrIyrValLeuLysLysGIy 661
Qy 873 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTTCAACATCAATACGTT 932
Db 662 GIuLysProIyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysIyrVal 681
Qy 933 GATGTGATACCAACGAATTGCTTAAAAAGTAGACAGCTTTTAAACGTTAGCGAACTTAC 992
Db 682 AspValAspThrAsnGIuLeuLeuLysSerGIuGIuLeuThrAlaSerGIuArgAsn 701
Qy 993 TTACACTTCAGAGATTTTATAGATCCCTCGTATAGGTAAGCTTAACCTTACAACAATTC 1052
Db 702 LeuAspPheAlaGAspLeuIyrAspProArgAspLysAlaLysLeuLysIyrAsnAsnLeu 721
Qy 1053 GATGCTTTGGTATATGAGCTATACCTTATCTGAAAGAGTAGAGATTAACAGATGAC 1112
Db 722 AsphAlaPheGIyIleMetAspIyrThrLeuThrgIySValGIuAspAsnHisAspAsp 741
Qy 1113 ACCAACCGTATCATTAACCGTTTATATAGGCAAGCAGCCGGAAGAGGAATGCTAGCTAC 1172
Db 742 ThrAsnArgIleIleThrValIyrMeGlyLysArgProGIuGIuAsnAlaSerIyr 761
Qy 1173 CATTTAGCT 1181
Db 762 HisLeuAla 764
RESULT 6
US-07-854-596B-47
; Sequence 47, Application US/07854596B

Patent No. 5434073
 GENERAL INFORMATION:
 APPLICANT: Dawson, Keith M
 APPLICANT: Hunter, Michael G
 APPLICANT: Caplewski, Lloyd G
 TITLE OF INVENTION: Proteins and nucleic acids
 NUMBER OF SEQUENCES: 73
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Dr. John J. McDonnell
 STREET: Ten South Wacker Drive, Suite 3000
 CITY: Chicago
 STATE: IL
 COUNTRY: USA
 ZIP: 60606
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/854,596B
 FILING DATE: 03-JUN-1992
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: McDonnell, John J
 REGISTRATION NUMBER: 26,949
 REFERENCE/DOCKET NUMBER: 92,337
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312-715-1000
 TELEFAX: 312-715-1234
 TELEX: 910-221-5317
 INFORMATION FOR SEQ ID NO: 47:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 483 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-07-854-596B-47

Alignment Scores:
 Pred. No.: 1,266-191 Length: 483
 Score: 1978.00 Matches: 401
 Percent Similarity: 82.28% Conservative: 17
 Best Local Similarity: 78.94% Mismatches: 52
 Query Match: 72.83% Indels: 38
 DB: 1 Gaps: 8

US-09-940-235-9 (1-1541) x US-07-854-596B-47 (1-483)

QY 33 ATAGTGTCTGATGATGCTACTAGATGCTCTTCTGTAATAATACAGCAATTGTTGTT 92
 DB 1 IleaaglyProglutripLeuaspargProSerValasnasSerGlnleuVal 20
 QY 93 AGCGTGTGTAAGTGTGAAGGGAAGATCAAGACATAGTCTTAATTTTGAATC 152
 DB 21 SerValaaglyThValGluGlyThranglnaSerleuysSphenegluVal 40
 QY 153 GATTAACATCAAGCCTGCTCATGAGAGAAAGACAGCAAGGCTTAAGTCAAAATCA 212
 DB 41 AspleuThrSerArgProalalHisgllyGlysthrGlnGlnGlyleuSerProlysSer 60
 QY 213 AAACCATTTGCTACTGATAGTGGCGGAGATGTCACATAACTTGAGAAAGCTTACTA 272
 DB 61 LysProPhealathAspSerGlyAlaMetProHislyleuGlyLysAlaAspleuLeu 80
 QY 273 AAGGCTATTCAAGAACATTTGATGCTAACGTCACAGTAAGAGAGACTTATGAGTTC 332
 DB 81 LysalalieglnGlnGlnleuIlealaaenValHisSerAsnaspPyrPhegluVal 100
 QY 333 ATTAATTTGCAAGGATGCAACCATTAAGTTCGAAACGCAAGGCTTACTTTGCTGAC 392
 DB 101 IleasPhealaserAspAlathrllethraPargAsnGlylysvaValYrPhealaaP 120

QY 393 AAAGATGTTGGTAACTTGCCGAGCAACCACTGTCACAGAAATTTTGTAAAGCGACAT 452
 DB 121 LysaspGlySerValThrleuProThrlnProValGlnGlnPheleuLeuSerGlyHis 140
 QY 453 GTGGCGTTAGACCATTAAGAAAGAAACAAATACAAACCAAGGAAATGTTGATGTG 512
 DB 141 ValArgValArgProGlyLysGlnYsProIleGlnasnglnAlaLysSerValAspVal 160
 QY 513 GAATPACTGTACAGTTTACTCCCTTAAGCCCTGAAGTACGATTTTCAACAGGCTCAAA 572
 DB 161 GluYrThrValGlnPheThrProleuAsnProAspAspPheArgProGlyLeuLys 180
 QY 573 GATPACTAGCTATTGAAGAACATAGTATGCTGACACCATCATCATCTCAAGAAATTA 632
 DB 181 AspThrLysleuLeuysThrleuAlaIleGlyAspThrIleThrSerGlnleuLeu 200
 QY 633 GCTCAAGCAACAAACATTTTAAACAAACCAACCCAGGCTTAACGATTTATGAACGTGAC 692
 DB 201 AlaglnalaglnSerIleleuAsnLysThrHisProGlyYrThrIleYrGluArgAsp 220
 QY 693 TCCCTCAATGCTCACTGACATGACATGATTTCCGTAACATTTTCAACATGATCAAGAG 752
 DB 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleuPrometAspGlnGlu 240
 QY 753 TTACTTACCGTGTAAATAATCGGGAACAGCTTAATGATCAATAATAAATCGTCTG 812
 DB 241 PheThrTrpHisValLysAsnArgGlnGlnAlaIleGlyIleAsnLysLysSerGlyLeu 260
 QY 813 AATGAAGAAATTAACAAACACTGACCTGATCTGAGAAATTAACGTCCTTAATAAAGG 872
 DB 261 AsnGlnGlnIleAsnSerThrAspLeuIleSerGlnLysYrYrValleuLysGly 280
 QY 873 GAAAAGCCGTAATGATCCCTTGTATCGGAGTCACTTGAACCTGTACCATCAATAGT 932
 DB 281 GluLysProGlyAspProPheAspArgSerHisleuLysleuPheThrIleYrYrVal 300
 QY 933 GATGCGATPACCAAGCAATGCTTAAGAGAGAGAGCTTAACAGCTAGCAAGCACTGAC 992
 DB 301 AspValAsnThrAsnGlnleuLeuLysSerGlnGlnleuThrAlaSerGlnArgAsn 320
 QY 993 TTAGACTTCAGAGATTTATACATCCCTGATAGAGGCTTAACACTTCAACAACTTC 1052
 DB 321 LeuAspPheArgAspLeuYrAspProArgAspLysAlaLysleuLeuYrAsnAsnLeu 340
 QY 1053 GATGCTTTTGGTATTAAGACTATTAACCTTAACCTGAGAAAGTAAATTAACAGATGAC 1112
 DB 341 AspAlaPheGlyIleleuMetAspYrThrleuThrGlyYrValGluAsnHisAspAsp 360
 QY 1113 ACCAACCGTATCATTAACCGTTTATATGCGCAAGCAACCGAAGAGAAATGCTAGCTAC 1172
 DB 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGlnGlnGlnAlaSerYr 380
 QY 1173 CATTTAGCTGTGTGTGCGCAGCGCAACAGATTGTACCCATAGCTGAGAAATGTTTGTAT 1232
 DB 381 HisleuAla-----TyrAspLysAspArgYrThrGlnGlnGlnAlaValTyrSer 398
 QY 1233 CATGCTGTGGGACTTCCTATATGCTGTGCGAAGAACTGGAGAAACCTTACCAAGCTGG 1292
 DB 399 TyrleuArgYrThrGlyThrProIleProAsnProAsnAspLysIleGlnGlyArg 418
 QY 1293 ATGATG---CTAATGTTACT-----TGCCGTGGAGAAAGGAGGGA 1331
 DB 419 ValValTyrThrAspCysThrGlnSerGlyGlnAsnleuLysCysGlnGlySerAsn 438
 QY 1332 CGCATCACTTGACACTTCTAAGAAATTAAGTCAACAGATCAGACACAAAGATCCTATAGA 1391
 DB 439 -----ValCysGlyGlnGlnGlnAsnLysCys-----Ile 447
 QY 1392 ATTTGAGACACCTTGAGCAAGAAAGATTAATGAGAAACCTGCTCAAGTCACTGACCA 1451
 DB 448 LeuGlySerAspGlyGlnLysAsn-----GlnCysVal---Thr 459
 QY 1452 GGCACAGCGCCGAGAGAGTGAAGTGTGAGAGGACACCTCTGTGACAGACCAATCGAC 1511

```

Db      460 GlyGluGly-----ThrProLysProGlnSerHis 469
Oy      1512 GGATCTGGCCCTTGCACCGATGTT 1535
Db      470 AsnAspGlyAspPheGluGluIle 477

RESULT 7
US-08-488-940-3
; Sequence 3, Application US/08488940
; Patent No. 5854049
; GENERAL INFORMATION:
; APPLICANT: Reed, Guy L.
; TITLE OF INVENTION: PLASMIN-RESISTANT STREPTOKINASE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,940
; FILING DATE: 09-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 05433/009001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 813 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-488-940-3

Alignment Scores:
Pred. No.: 2,4e-190 Length: 813
Score: 1966.50 Matches: 382
Percent Similarity: 99.74% Conservative: 0
Best Local Similarity: 99.74% Mismatches: 0
Query Match: 72.40% Indels: 1
DB: 1 Gaps: 1

US-09-940-235-9 (1-1541) x US-08-488-940-3 (1-813)
Oy      33 ATAGCTGCTCTGAATGCTACTAGATCGCTTCTGTAATAATAACGCAATGTTGTT 92
Db      1 IleAlaGlyProGluThrLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
Oy      93 AGCGTGTCTGTAAGTCTTGAAGGGAGCAATCAAGACATTAGCTTAAATTTTGAATC 152
Db      21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPheGluIle 40
Oy      153 GATCTTAACATCAAGACCTGTCTCATAGAGAAAGACGAGAGAGCTTAAGTCCAAATCA 212
Db      41 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 60
Oy      213 AAACCATTTGTACTAGATAGTGGCGGATGTCACTAACTTGAAGAAGCTGACTTACTA 272
Db      61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlnLysAlaAspLeuLeu 80

```

```

Oy      273 AAGCTATTCAAGAAACATTGATGCTTAAGCTCCAGATTAACGAGCTACTTGAAGTTC 332
Db      81 LysAlaIleGlnGluGlnLeuIleAlaSerValHisSerAsnAspArgPheGluVal 100
Oy      333 ATTGATTTTGGAGCGATGCAACCATTACTGATCGAAAACGGCAAGGTCTTACTTGCAC 392
Db      101 IleAspPheAlaSerAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
Oy      393 AAAGATGTTGGTGAACCTTGGCCGACCCAGCTGTCCAAAGAAATTTTCTTAAGGAGAT 452
Db      121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
Oy      453 GTGGCGCTTAGACCATTTAAAGAAAACCAATACAAAACCAAGCAAAATCTGTGATGTC 512
Db      141 ValArgValArg---TyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 159
Oy      513 GAATATACGTGATACATTACTCCCTTAAACCCGTATGACGATTTCAACGAGTCTCAAA 572
Db      160 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 179
Oy      573 GATCTTAAGCTATTGAAAACACTAGCTATCGGTGACACATCACTCAAGAAATTACTA 632
Db      180 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnLeuLeu 199
Oy      633 GCTCAAGCAACAAGCATTTTAAACAAAACACACGAGCTATGATTTTGAAGTGCAC 692
Db      200 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 219
Oy      693 TCCTCAATCGTCACTGATGACAATGACATTTTCCGTACGATTTTACCAATGATCAAG 752
Db      220 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 239
Oy      753 TTACTTACCGGTGTTAAAAATCGGAAACAGCTTATAGATCAATAAAAATCTGCTCG 812
Db      240 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysSerGlyLeu 259
Oy      813 AATCAAGAAATTAACAACACTGACCTGATCTGTGAAATTTACGTCCTTAAAAAGCG 872
Db      260 AsnGluGlnIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysGly 279
Oy      873 GAAAAGCGATGATGCCCTTGTGATCGAGCTCACTTGAACCTGACCATCAATGATCTC 932
Db      280 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheMetIleLysTyrVal 299
Oy      933 GATGTGATACCAACGAATGCTTAAAAAGTAGAGAGCTTTAACAGCTAGCAACGTAC 992
Db      300 AspValAspThrAsnGlnLeuLeuLysSerGluGlnLeuThrAlaSerGluArgAsn 319
Oy      993 TTAGACTTCAGAGATTTATACGATCTCGTATAGAGCTTAACCTACTTACACAAATCTC 1052
Db      320 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAspAsnLeu 339
Oy      1053 GAGCTTGTGTATTAATGACCTATACCTTAACGTGAAGAAGTAGAGATTAATCAAGATGAC 1112
Db      340 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyValGluAspAsnHisAspAsp 359
Oy      1113 ACCAACCGTATCAATACCGTTTATATGAGCAAGGACCCGAAAGAGAGAAATGTAGCTAC 1172
Db      360 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGluAsnAlaSerTyr 379
Oy      1173 CATTTAGCT 1181
Db      380 HisLeuAla 382

RESULT 8
US-08-488-940-1
; Sequence 1, Application US/08488940
; Patent No. 5854049
; GENERAL INFORMATION:
; APPLICANT: Reed, Guy L.
; TITLE OF INVENTION: PLASMIN-RESISTANT STREPTOKINASE
; NUMBER OF SEQUENCES: 20

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,940
FILING DATE: 09-JUN-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Frazer, Janis K.
REGISTRATION NUMBER: 34,819
REFERENCE/DOCKET NUMBER: 05433/009001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ. ID NO.: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1194 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-488-940-1

Alignment Scores:		
Pred. No.:	2.9e-190	Length: 119
Score:	1966.50	Matches: 382
Percent Similarity:	99.74%	Conservative: 0
Best Local Similarity:	99.74%	Mismatches: 0
Query Match:	72.40%	Indels: 1
DB:	1	Gaps: 1

Oy	33	TTAGCTGGGCTCGAATGGCTACTAGATCGCCCTTGTAATTAACAAGCAATGGTGTGTT	92
Db	382	lleaaglyprrglintrip.leuenuapargproSerValaasmnserglnleuValVal	401
Oy	93	AGCGTGTGCTACTGTTCGAGGGAGCAATCAAGACTTAGTCTTAAATTTTGAATC	152
Db	402	SerValaaglyThnValgluThnraenglnasprlIeSerleuysphreheglnlle	421
Oy	153	GATCTAACATCAAGACTTGCTCTCATGAGGAGAAAGACAGAGCAAGCTTAACTCCAAATCA	212
Db	422	AspIeuThnSerArgProalahlsglyglYlyThnrglnlgnlYleuSerProlySer	441
Oy	213	AAACCATTTGCTACTGATATGTGGCGGAGATGTCATATAACTTGAAGAAAGCTGACTTACTA	272
Db	442	lysProphelaatThraAspserglYalaIaetSerhIaIySleuGluluySbaIaasprleuIeu	461
Oy	273	AAGGCTATTCAAGAACAATTGATCGGTAAACGTACACAGTAAAGACGCTACTTTGAGGTC	332
Db	462	lysbaIaIleglnglnlgnlneulIealaasrValaHnISerasmnspasprlyrPhnglulVal	481
Oy	333	ATTGATTTTTCGAAGCGATGCAACCATTTACTGATCGAAACGGCAAGGTCTACTTTGGCTGAC	392
Db	482	IleaasphneIaIaSeraspIaIaThrlIeThraAspargasnclYlySvalYlyrPhneIaIaAsp	501
Oy	393	AAAGATGGTTCGGTAACTTTCGCGAGCCCAACTGTGCCAAGATTTTGTCTAACCGGACAT	452
Db	502	lysaaSprglySerValThrIeuProTrngInProValglnglnIuPhleuIeuSerIyHnIS	521
Oy	453	GTGCGCGTTAAGCAATATTAAGAAAAACAATTCAGAAACAAGCGAAATCTGTGTGATGTG	512

```

Db      522 ValaGValaGg---TyltlyGslwysProIleGlnaGlnAlaIlySerValaAspVal 540
QY      513 GAATTAATATGTAACAGTTACTTCCCTTAAACCCCTGATGACGATTTAGACCAAGTCTCAA 572
Db      541 GltvYrThrValaGlnPheThrProLeuasnProaspaspPheArgProGlyLeuIlys 560
QY      573 GATACATAAGCATTTGAAAACAAGTACTGCTGCTGTCGACACATCAATCAATCCAGAAATTACCA 632
Db      561 AspThrIlySleuLeuIysThrIleuAlaIleGlyAspThrIleThrSerGlnGluLeu 580
QY      633 GCTCAAGCACAAAGCATTTTAAACAAAACCAACCCAGGCTATACGATTTATGAACGTGAC 692
Db      581 AlaGlnAlaGlnSerIleLeuAsnIlyAsnHisProGlyIlyThrIleTyrGlnArgAsp 600
QY      693 TCTCTCAATGCTGACTCATGACAAATGACATTTTCCGTACGATTTTACCATGGATCAAGAG 752
Db      601 SerSerIleValaIleThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 620
QY      753 TTTACTTAACGGTGTAAAAATCCGGGAACAACCTTATAGACATCAATAAAAAATTCGGTGTG 812
Db      621 PheThrIlyrValaValIyAsnArgGlnGlnAlaIlyrArgIleAsnIlySlySerIlyLeu 640
QY      813 AATGAGAAGAAATAAACAACTGACCTGATCTCTGAGAAATATTACGTCTTAAAAAAGG 872
Db      641 AsnGlnGlnIleAsnAsnThrAspLeuIleSerGlnIlyrIlyrValLeuIlySly 660
QY      873 GAAAAAGCGTATGATCCCTTTGATGCGAGTCACATTTGAAACGTGTCCATCAATACATGTT 932
Db      661 GltIysProIlyrAspProPheAspArgSerHisIleuIlySleuPheThrIleTyrVal 680
QY      933 GATGTCCGATTAACAAGAAATTTGCTTAAAAAGTAGACGCTCTTAAACGCTAGGAACGTAC 992
Db      681 AspValAspThrAsnGlnLeuLeuIlySerGlnGlnLeuLeuThrAlaSerGlnArgAsn 700
QY      993 TTAAGACTTCAGAGATTTATATACGATCTCTGTATAGAGCTTAAAGCTACTTACAAACATCTC 1052
Db      701 LeuAspPheArgAspLeuIlyrAspProArgAspIlySalaIlySleuIlyrAsnAsnLeu 720
QY      1053 GATGCTTTGGTATATATGACTATATACCTTAACTGAAAAAGTAGAGATATACGATGAC 1112
Db      721 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyIysValGlnAspAsnHisAspAsp 740
QY      1113 ACCAACCCTATCATTAACCGTTTATATGGGCAAGGACCCGAAAGAGAAATCTTAGCTAC 1172
Db      741 ThrAsnArgIleIleThrValIlyMetGlyIlyAsArgProGlnGlyGlnAsnAlaSerTyr 760
QY      1173 CATTTAGCT 1181
Db      761 HisLeuAla 763

RESULT 9
US-07-854-596B-26
: Sequence 26, Application US/07854596B
: Patent No. 5434073
:
: GENERAL INFORMATION:
:   APPLICANT: Dawson, Keith M
:   APPLICANT: Hunter, Michael G
:   APPLICANT: Czaplowski, Lloyd G
:   TITLE OF INVENTION: Proteins and nucleic acids
:   NUMBER OF SEQUENCES: 73
:   CORRESPONDENCE ADDRESS:
:     ADDRESSER: Dr. John J. McDonnell
:     STREET: Ten South Wacker Drive, Suite 3000
:     CITY: Chicago
:     STATE: IL
:   COUNTRY: USA
:   ZIP: 60606
:
: COMPUTER READABLE FORM:
:   MEDIUM TYPE: Floppy disk
:   COMPUTER: IBM PC compatible
:   OPERATING SYSTEM: PC-DOS/MS-DOS
:   SOFTWARE: PatentIn Release #1.0, Version #1.25
:
: CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: US/07/854,596B
FILING DATE: 03-JUN-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: McDowell, John J
REGISTRATION NUMBER: 26,949
REFERENCE/DOCKET NUMBER: 92,337
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
TELEX: 910-221-5317
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 415 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-854-596B-26

Alignment Scores:

Pred. No.:	3,89e-190	Length:	415
Score:	1963.00	Matches:	379
Percent Similarity:	98.96%	Conservative:	1
Best Local Similarity:	98.70%	Mismatches:	4
Query Match:	72.28%	Indels:	0
DB:	1	Gaps:	0

US-09-940-235-9 (1-1541) x US-07-854-596B-26 (1-415)

```
QY 30 ATGATAGCTGCTCTGTAATGCGCTACTAGATCGCTCTGTGTAATTAACGCAATGCTT 89
   |||
Db 1 MetIleAglYProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20

QY 90 GTTACGCTTGTGTAAGTCTGTAAGGGAAGCAATGCAATGCTTAAATTTTGTGA 149
   |||
Db 21 ValSerValAglYThrValGluGlnYThrAsnGlnAspIleSerLeuYSerPheGln 40

QY 150 ATCATGTAACATGACGACTGCTCATGAGAGAAAGACGACGAGCTTAAGTCCAAA 209
   |||
Db 41 IleAspLeuThrSerArgProIleAsnIleGlyGlnGlnGlnGlnGlnGlnGlnGlnGln 60

QY 210 TCAGAACCATTTGCTACTGATAGTGGCGCATGTGCATTAACCTTGAAGAACTGACTTA 269
   |||
Db 61 SerYSerProPheAlaThrAspSerGlyAlaMetProHISylSerGlnYSerIleAspLeu 80

QY 270 CTAAAGGCTATTCAAGCAATTTGATCGCTTAACGTCCACAGTACGACGACTTGTGAG 329
   |||
Db 81 LeuYsAlaIleGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 100

QY 330 GTGATGATTTTGAAGCGATGCAACCATTTACTGATGAAACGGCAAGCTTACTTGTCT 389
   |||
Db 101 ValIleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValYrPheAla 120

QY 390 GACAAAGATGGTTCGTAACCTTCCGACCAACCTGTCCAAAGATTTTGTGAAGCGGA 449
   |||
Db 121 AspYsAspGlySerValThrLeuProThrGlnProValGlnGlnPheLeuLeuSerGly 140

QY 450 CATGTGCGCTTGAACCATTAATAAGAAACCAATATACAAACCAAGCGAAATCTGTGAT 509
   |||
Db 141 HisValArgValAglProYrYrGlnYrSerProIleGlnAsnGlnAlaLysSerValAsp 160

QY 510 GTGGAATATCTGTACAGTTTACTCCCTTAAACCTTGAACGATTTGACACAGGCTTC 569
   |||
Db 161 ValGlnYrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeu 180

QY 570 AAGATCTAGCATTTGAAACCTAGCTATCGGTGACACATACATCTCAAGATTA 629
   |||
Db 181 LysAspThrYrYsLeuLeuYrThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeu 200

QY 630 CTAGCTCAAGCAACAGATTTTAAACAAACCAACCGAGCTATACGATTTATGAGCT 689
   |||
Db 201 LeuAlaGlnAlaGlnSerIleLeuAsnYrThrHisProGlyYrThrIleYrGlnYrG 220
```

```
QY 690 GACTCCTCAATCGCTACCTACATGACAAATGACATTTCCGTACGATTTTACCAATGATCAA 749
   |||
Db 221 AspSerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGln 240

QY 750 GACTTACTTACCGGTGTTAAATCCGGAAACAGCTTATAGATCATTAATAAATCTGCT 809
   |||
Db 241 GluPheThrYrHisValYrYsAsnArgGlnGlnAlaYrGlnIleAsnYrYsSerGly 260

QY 810 CTGATGAAGAATAAACAACACTGACCTGATCTGAGAAATPTAGCTTCAAAAA 869
   |||
Db 261 LeuAsnGlnGlnIleAsnMetThrAspLeuIleSerGlnYrYrYrValLeuYrYs 280

QY 870 GGGGAAAAGCCGTATGATCCCTTGTGATCGCATCTTGAACCTTGAACCTTCAACAAATAC 929
   |||
Db 281 GlyGlnYrProYrYrAspProPheAspArgSerHisLeuYrYsPheThrIleYrYr 300

QY 930 GTTGAATGTCATACCAACGATTTGCTTAAAGTGAAGAGAGCTTTAAACAGTACGAAAGT 989
   |||
Db 301 ValAspValAsnThrAsnGlnLeuLeuYrSerGlnGlnLeuLeuThrAlaSerGlnArg 320

QY 990 AACTAGACTTCGAGATTTATACGATCCTCGTATAGGCTAACTTCAACAAAT 1049
   |||
Db 321 AsnLeuAspPheArgAspLeuYrYrAspProArgAspYrYrYrYrYrYrYrYrYrYrYr 340

QY 1050 CTCGATGCTTTTGTATTTATGACTATACCTTAACCTGAAGAAAGTGAAGATTAACGAT 1109
   |||
Db 341 LeuAspAlaPheGlyIleMetAspYrYrThrLeuThrGlyYrYrValGlnAspAsnHisAsp 360

QY 1110 GACACCAACCGTATCATTAACCGTTTATATGAGGACGACCGGAAAGAGAAATGCTAGC 1169
   |||
Db 361 AspThrAsnArgIleIleThrValYrMetGlyYrYrYrYrYrYrYrYrYrYrYrYrYrYr 380

QY 1170 TACGATTTAGCT 1181
   |||
Db 381 TyrHisLeuAla 384
```

RESULT 10

US-07-854-596B-19

Sequence 19, Application US/07854596B

Patent No. 5434073

GENERAL INFORMATION:

APPLICANT: Dawson, Keith M

APPLICANT: Hunter, Michael G

APPLICANT: Czaplowski, Lloyd G

TITLE OF INVENTION: Proteins and nucleic acids

NUMBER OF SEQUENCES: 73

CORRESPONDENCE ADDRESSES:

ADDRESSER: Dr. John J. McDowell

STREET: Ten South Wacker Drive, Suite 3000

CITY: Chicago

STATE: IL

COUNTRY: USA

ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/854,596B

FILING DATE: 03-JUN-1992

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: McDowell, John J

REGISTRATION NUMBER: 26,949

REFERENCE/DOCKET NUMBER: 92,337

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-715-1234

TELEX: 910-221-5317

INFORMATION FOR SEQ ID NO: 19:

SEQUENCE CHARACTERISTICS:

LENGTH: 415 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-854-596B-19

Alignment Scores:

Pred. No.:	1.28e-189	Length:	435
Score:	1958.00	Matches:	378
Percent Similarity:	98.96%	Conservative:	1
Best Local Similarity:	98.69%	Mismatches:	4
Query Match:	72.09%	Indels:	0
DB:	1	Gaps:	0

US-09-940-235-9 (1-1541) x US-07-854-596B-19 (1-435)

```

QY 33 ATAGTGTGCTCCGAATGGTACTAGATCGTCTTCTGTAAATPACAGCCAAATTGTTGTT 92
DB 22 IleaGlyProGluThrLeuAspArgProSerValAsnAsnSerGlnLeuVal 41
QY 93 AGCGTTGCTGTAAGGAGGAGCAATCAAGACATTAATGTTAAATTTTGAATC 152
DB 42 SerValaGlyThrValGluGlnThrAsnGlnAspIleSerLeuysPhePheGluIle 61
QY 153 GATGTACATCAAGACCTGCTCATGAGGAGAAAGACAGCAAGGCTTAAGTCCAAATCA 212
DB 62 AspLeuThrSerArgProAlaHisGlyGlySerThrGluGlnGlyLeuSerProlySer 81
QY 213 AAACCATTTGCTACTGATGAGGCGGATGTCAATAACTTGAGAACTGACTTA 272
DB 82 LysProPheAlaThrAspSerGlyAlaMetProHisbyleuGlnIulysAlaAspLeu 101
QY 273 AAGGTATTCAAGAACATTTGATCGTACGTCACAGTAAGACGACTTGTAGGTC 332
DB 102 LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspArgPyrRheGluVal 121
QY 333 ATTGATTTTGGCAAGGAGATGCAACATTAATGTCGAAACGCAAGGCTTACTTTGTCAC 392
DB 122 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyValIleValIleAsp 141
QY 393 AAAGTGTGCTGGTAACCTTGCCGACCCAACTGTGCCAAGAAATTTTGTAAAGCGACAT 452
DB 142 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuSerGlyHis 161
QY 453 GTGGCGGTTAGCCATTTAAAGAAAAACCAATACAAAACCAAGCCAAATCTGTATGTCG 512
DB 162 ValArgValArgProGlyGlySerProIleGlnAsnGlnAlaLysSerValAspVal 181
QY 513 GAATATCTGTACAGTTTACTCCCTTAAACCTGATGACATTTTTCAGACAGGTCGAA 572
DB 182 GluThrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 201
QY 573 GATACATAAGCTATTGAAAACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTA 632
DB 202 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 221
QY 633 GCTCAAGCAAGAGATTTTAAACAAAACCAAGGCTATACGATTTATGAACTGAC 692
DB 222 AlaGlnAlaGlnSerIleLeuAsnLysThrHisAspGlyThrIleThrGluAspArg 241
QY 693 TCCCTCAATCGTCACTCATGACATGACATTTTCCGTAAGATTTTTCAGATGATCAAGAG 752
DB 242 SerSerIleValThrHisAspAsnAspIlePheAsnGlnThrIleLeuProMetAspGlnGlu 261
QY 753 TTTACTTAACGTTGTTAAAAATCGGAGCAAGCTTATAGGATCAATAAAAATCTGGTCTG 812
DB 262 PheThrThrHisValLysAsnArgGluGlnAlaThrGlnIleAsnLysLysSerGlyLeu 281
QY 813 AATGAAGAAATTAACAACACTGATCTGTGAGAAATATTACTCTTAAAAAGG 872
DB 282 AsnGluGluIleAsnAsnThrAspLeuIleSerGlyLysThrGlyValLeuLysGly 301
QY 873 GAAAGCGGTATGATCCCTTGTATGCGACGACTTGAACCTGTACCATCAATCTGCT 932

```

```

DB 302 GluLysProGlyAspProPheAspArgSerHisLeuLysLeuPheThrIleLysThrVal 321
QY 933 GATGTGATPACCAAGCAATTTGCTAAAGAGACAGCTTTACAGCTACGAAACGTAAC 992
DB 322 AspValAsnThrAsnGlnLeuLysSerGlnGlnLeuThrAlaSerGluArgAsn 341
QY 993 TTAGACTCGAGATTTATACATCCCTCGATAGGCTAAACTACTCTCAACAACTCTC 1052
DB 342 LeuAspPheArgAspLeuThrAspProArgAspValAlaLysLeuLeuThrAsnAsnLeu 361
QY 1053 GATGCTTTGGATTATGACATATACCTTAACCTGAAAGATAGAGATAATCAAGTAC 1112
DB 362 AspAlaPheGlyIleMetAspThrThrLeuThrGlyValGluAspAsnHisAspAsp 381
QY 1113 ACCAAGCTATCATACCGTTTATATGGGACAGACCCGAAAGAGAAATGCTAGTAC 1172
DB 382 ThrAsnArgIleIleThrValThrMetGlyLysArgProGluGluAsnAlaSerThr 401
QY 1173 CATTTAGCT 1181
DB 402 HisLeuAla 404

RESULT 11
US-07-854-596B-15
; Sequence 15, Application US/07854596B
; Patent No. 5434073
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith M
; APPLICANT: Hunter, Michael G
; APPLICANT: Czaplowski, Lloyd G
; TITLE OF INVENTION: Proteins and nucleic acids
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESSES:
; ADDRESSER: Dr. John J. McDonnell
; STREET: Ten South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/854,596B
; FILING DATE: 03-JUN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,337
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 440 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-854-596B-15

Alignment Scores:
Pred. No.: 1.29e-189 Length: 440
Score: 1958.00 Matches: 378
Percent Similarity: 98.96% Conservative: 1
Best Local Similarity: 98.69% Mismatches: 4
Query Match: 72.09% Indels: 0
DB: 1 Gaps: 0

US-09-940-235-9 (1-1541) x US-07-854-596B-15 (1-440)

```

```

QY 33 ATAGCTGCTCTGAATGGCTACTAGATCGTCTTCTGTAAATTAACAGCCAAATGGTGTGTT 92
DB 27 ILeAlaGlyProGluThrLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal 46
QY 93 AGCGTGTCTGTACTGTGGAGGGAAGATCAAGACATTAGTCTTAAATTTTGAATC 152
DB 47 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 66
QY 153 GATCTTAACATCAAGCCTGCTCATGAGGAGAAAGACAGAGAGCTTAAGTCCAAATCA 212
DB 67 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 86
QY 213 AAACCAATTTGCTACTGATAGTGGCGGATGTCAATTAACCTTGAGAAAGCTGACTTACTA 272
DB 87 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGluValAlaAspLeuLeu 106
QY 273 AAGGCTATTCAAGAACAAATTTGATCGCTAACGTCACAGTACAGACGACTTGTAGAGTC 332
DB 107 LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspArgTyrPheGluVal 126
QY 333 ATTTGATTTTGAAGCGATGCAACCATTAAGTCAAGGAGGAGGTCTTCTGCTGAC 392
DB 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyValTyrPheAlaAsp 146
QY 393 AAAGATGCTTCGGTAACTTTGCGGACCCAACTGTCACAAATTTTGTAAAGCGGACAT 452
DB 147 LysAspGlySerValThrLeuProThrGlnProValGlnGluLeuLeuSerGlyHis 166
QY 453 GTGCGGCTTGAACCATATATAAGAAAAACCAATACAAACAGCGAAATGTTGATGTG 512
DB 167 ValArgValArgProTyrLysGlyLysProIleGlnAsnGlnAlaLysSerValAspVal 186
QY 513 GAATATACCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTCAGACGAGCTTCAA 572
DB 187 GlnTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 206
QY 573 GATCTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATGCAATCTCAAGAAATTA 632
DB 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 226
QY 633 GCTCAAGCAACAAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTTGAACGTCAC 692
DB 227 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGlnAspArg 246
QY 693 TCTTCATGTGCTACATGACATGACATGATTTTCCGATCGATTTTCAATGATGATCAAG 752
DB 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
QY 753 TTTACTTACCGGTAAAAAATGGGAAACAAAGCTTAATAGATCAATAAAAATCTGGTCTG 812
DB 267 PheThrTyrHisValLysAsnAspGlnGlnAlaTyrGlnIleAsnLysLysSerGlyLeu 286
QY 813 AATGAAGAAATTAACACACATGACTGATCTCTGAGAAATATTAACGCTTAAAAAAAGG 872
DB 287 AsnGluGluIleAsnAsnThrAspLeuIleSerGlnLysTyrTyrValLeuLysLysGly 306
QY 873 GAAAAGCCGTATGATCCCTTGTATGCGACGCTGATGAAACTGTTCACATCAATACGTT 932
DB 307 GlnLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326
QY 933 GATTCGCAATCAACGAATTTGCTAAAAAGTGAAGACTCTTAAACAGCTACGGAAGTAA 992
DB 327 AspValAlaGlyThrAsnGluLeuLeuLysSerGlnGlnLeuThrAlaSerGluAsnAsp 346
QY 993 TTAGACTTCAGAGATTTATAGCATCCTCGTGAATAAGGCTAACTACTACAAATTC 1052
DB 347 LeuAspPheAspAspLeuTyrAspProAspAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
QY 1053 GATGCTTTTGTATTATGACTATACCTTAACTGAAAAGTAGAGGATTAATCAAGTAC 1112
DB 367 AspAlaPheGlyIleMetAspTyrThrIleuThrGlyLysValGlnLysAsnHisAspAsp 386

```

```

QY 1113 ACCAAGCTATCATTAACCGTTTATATGGGCAAGGACCCGGAAGAGAGAAATGCTAGTAC 1172
DB 387 ThrAsnArgIleIleThrValTyrMetGlyAsnArgProGluGlnAsnAsnSerTyr 406
QY 1173 CATTTAGCT 1181
DB 407 HisLeuAla 409

RESULT 12
US-07-854-596B-43
; Sequence 43, Application US/07854596B
; Patent No. 5434073
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith M
; APPLICANT: Hunter, Michael G
; APPLICANT: Czaplowski, Lloyd G
; TITLE OF INVENTION: Proteins and nucleic acids
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. John J. McDonnell
; STREET: Ten South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/854,596B
; FILING DATE: 03-JUN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,337
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEEX: 910-221-5317
; INFORMATION FOR SRO ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 483 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-854-596B-43

Alignment Scores:
Pred. No.: 1,35e-189 Length: 483
Score: 1958.00 Matches: 378
Percent Similarity: 98.96% Conservative: 1
Best Local Similarity: 98.69% Mismatches: 4
Query Match: 72.09% Indels: 0
Gaps: 0

US-09-940-235-9 (1-1541) x US-07-854-596B-43 (1-483)
QY 33 ATAGCTGCTCTGAATGGCTACTAGATCGTCTTCTGTAAATTAACAGCCAAATGGTGTGTT 92
DB 27 ILeAlaGlyProGluThrLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal 89
QY 93 AGCGTGTCTGTACTGTGGAGGGAAGATCAAGACATTAGTCTTAAATTTTGAATC 152
DB 47 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 109
QY 153 GATCTTAACATCAAGCCTGCTCATGAGGAGAAAGACAGAGAGCTTAAGTCCAAATCA 212
DB 67 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 129
QY 213 AAACCAATTTGCTACTGATAGTGGCGGATGTCAATTAACCTTGAGAAAGCTGACTTACTA 272

```

```

Db 130 LysProPheAlaThrXAspSerGlyAlaMetProHleuValLeuGluValLysValAspLeuVal 149
QY 273 AAGGTAATTCAGAGCAATTCGCTAAGCTCCAGATGACAGCACTTCTTGAGGTC 332
Db 150 LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspArgPyrPheGlnVal 169
QY 333 ATTGATTTTGCAGGCGATGCAACCATTCATGATCGAAAGCGGCAAGGCTCTTGGCTGAC 392
Db 170 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValIlePheAlaAsp 189
QY 393 AAAAGATGGTTCGGTAACTTTCGCCAGCCAACTGTCGCAAGATTTTTCAGCGGACAT 452
Db 190 LysAspGlySerValThrLeuProThrGlnProValGlnGlnPheLeuLeuSerGlyHis 209
QY 453 GTGCGCGTTCAGACATATTAAGAAAAACCAATTCAAAACCAAGCGGAATTCGTGATGTC 512
Db 210 ValArgValArgProGlyLysGlyLysProIleGlnAsnGlnAlaLysSerValAspVal 229
QY 513 GAATATACCTGTACAGTTTACCTCCCTTAAACCTGATGAGATTCGAGACAGGTCGAA 572
Db 220 GlnIleThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 249
QY 573 GATATTAAGCTATTGAAAACTAGCTATCGGTGACACCATGACATCTCAAGATTAATA 632
Db 250 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeuLeu 269
QY 633 GCTCAAGCAACAAGCATTTTAAACAACAAACCAAGGCTATTCGATTTATGAACTGAC 692
Db 270 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyLysThrIleLysGluArgAsp 289
QY 693 TCCTCAATCGTCACTCATGACATGACATTTCCGATGAGATTTTACCATGATGCAAGG 752
Db 290 SerSerIleValAlaThrHisAspAsnAspIlePheAspGlyThrIleLeuProLeuAspGlnGln 309
QY 753 TTTACTTACCTCGTTTAAATAATCGGGAACAAGCTTATAGGATCAATTAATAATCGTCTG 812
Db 310 PheThrIleThrHisValLysAsnAspArgGlnGlnAlaIleGlyGlnIleAsnLysLysSerGlyLeu 329
QY 813 AATGAAGAATTAACAACAACCTGATCTGTGAGAAATATTAATTCCTTAAATAAGG 872
Db 330 AsnGlnGlnIleAsnAsnThrAspLeuIleSerGlnLysLysLysLysValLeuLysGly 349
QY 873 GAAAAGCCGATGATCCCTTATGATGCAATCCTTGAACCTGTCGACCAATTCGCT 932
Db 350 GlnLysProIleLysAspProPheAspArgSerHisLeuLysLeuPheThrIleLysIleVal 369
QY 933 GATGTCGATACCAAGCAATTCGTAATAAAGTGACAGCTCTTAAACAGCTAGCGCAACTAC 992
Db 370 AspValAlaThrHisGlnLeuLeuLysSerGlnGlnLeuThrAlaSerGlnAlaGln 389
QY 993 TTGAGCTTCAGAGATTTATGATGATCTCGTATTAAGGCTTAAACTACTTACCAACTTC 1052
Db 390 LeuAspPheArgAspLeuLysAspProArgAspLysAlaLysLeuLeuLysThrAsnAlaLeu 409
QY 1053 GATGCTTTTGGTATTAATGACATTAATCTTGAATAAGTAAAGATTAATCAAGATGAC 1112
Db 410 AspAlaAspGlyLysLeuAspPyrThrLeuThrGlyLysValGlnAspAsnHisAspAsp 429
QY 1113 ACCAAGCTATACATTAACGTTTATATGCGCAAGCGACCGCAAGAGAGATGCTAGCTAC 1172
Db 430 ThrAsnAlaGlnIleIleThrValIleThrGlyLysArgProGlnGlnGlnAlaSerIle 449
QY 1173 CATTTAGCT 1181
Db 450 HisLeuAla 452

```

RESULT 13
 US-07-854-596B-28
 ; Sequence 28, Application US/07854596B
 ; Patent No. 5434073
 ; GENERAL INFORMATION:
 ; APPLICANT: Dawson, Keith M

```

; APPLICANT: Hunter, Michael G
; APPLICANT: Czaplowski, Lloyd G
; TITLE OF INVENTION: Proteins and nucleic acids
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Dr. John J. McDonnell
; STREET: Ten South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/854,596B
; FILING DATE: 03-JUN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,337
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 499 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-854-596B-28

```

Alignment Scores:

```

Pred. No.: 137e-189 Length: 499
Score: 1958.00 Matches: 378
Percent Similarity: 98.96% Conservative: 1
Best Local Similarity: 98.69% Mismatches: 4
Query Match: 72.09% Indels: 0
DB: 1 Gaps: 0

```

US-09-940-235-9 (1-1541) x US-07-854-596B-28 (1-499)

```

QY 33 ATAGCTGCTCCGTAATGCTACTAGATCGCTCTTGTGAATAATACAGCCAAATGGTTGTT 92
Db 86 IleAlaGlyProGlnThrLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 105
QY 93 AGCGTTCGTGATCTGTTGAGGCGACGATCAAGACATTAAGCTTAAATTTTGAATC 152
Db 106 SerValAlaGlyThrValGlnGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 125
QY 153 GATCTAACATCAGACCTGCTCATGAGGAAAGACAGAGAGGCTTAACTCAACAAATCA 212
Db 126 AspLeuThrSerArgProAlaHisGlyGlyLysThrGlnGlnGlnLysSerProLysSer 145
QY 213 AAACATTTTCTCATAGTAGTGCGGAGATGTCATTAATCTTGAAGAGCTGACTTAATA 272
Db 146 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGlnLysValAspLeuLeu 165
QY 273 AAGGCTATTCAGAACAAATTCATGCTGTCACAGTACAGACGATCTTGAAGTC 332
Db 166 LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspArgPyrPheGlnVal 185
QY 333 ATTGATTTTGCAGGCGATGCAACCATTCATGATCGAAAGCGGCAAGGCTCTTGGCTGAC 392
Db 186 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValIlePheAlaAsp 205
QY 393 AAGAGATGTTCCGTTAATCTGCGGACCAACCTGTCGCAAGATTTTGTAAAGCGGACAT 452
Db 206 LysAspGlySerValThrLeuProThrGlnProValGlnGlnGlnPheLeuLeuSerGlyHis 225

```

QY 453 GTGCGGTTAGACCATATTAAGAAAAACAATACAAACCAAGCAATCTGTGATGTG 512
 Db 226 ValArgValArgProTyrIleGluGlyProIleGlnAsnGlnIleAlaGlySerValAspVal 245
 QY 513 GAATATACCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTCAGACCAAGCTTCAAA 572
 Db 246 GluTyrThrValGlnPheThrProIleAsnProAspAspPheArgProGlyLeuLys 265
 QY 573 GATATACGCTATTTGAAAAACATAGCTATCGTGACCACTGACATCTCAAGAAATTA 632
 Db 266 AspThrIleValLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 285
 QY 633 GCTCAAGCAAAAGCAATTTTAAACAAACCAACCAAGCTATACGATTTTGAACGTGAC 692
 Db 286 AlaGlnAlaGlnSerIleLeuLeuLysThrIleHisProGlyTyrThrIleTyrGlnIleAsp 305
 QY 693 TCCTCAATGCTGACATGACATGACATTTTCCGTAACGATTTTACCAATGATCAAG 752
 Db 306 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 325
 QY 753 TTTACTTACCGGTTTAAATGCGGAAACAAGCTTATAGATCAATATTAATTAATCTGCTG 812
 Db 326 PheIleTyrHisValIleValIleAsnArgGlnGlnAlaTyrGlnIleAsnLysSerGlyLeu 345
 QY 813 AATGAAGAATTAACCAACCACTGACCTGATCTCTGAGAAATATTAACGCTTAAACCAAGG 872
 Db 346 AsnGlnGlnIleAsnAsnThrAspLeuIleSerGlnLysTyrTyrValIleuLysGly 365
 QY 873 GAAAACCGTATGATCCCTTGAATGCGAGTCACTGAAACCTGTCAACATCAATAGCT 932
 Db 366 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 385
 QY 933 GATGTCGATACCAAGCAATTTGCTTAAACAGTACCTTTACACTGACCAAGCTAAC 992
 Db 386 AspValAsnThrAsnGlnLeuLeuLysSerGlnGlnIleuLeuThrAlaSerGlnIleAsn 405
 QY 993 TTAAGCTTCAGAGATTTATACGATCTCGATTAAGGCTTAACTACTCTTCAACCAATCTC 1052
 Db 406 LeuAspPheAspAspLeuTyrAspProAspArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 425
 QY 1053 GATGCTTTTGTATTTATGACATTAACCTTAACCTGAAAAGTAGAGGATTAACGATGAC 1112
 Db 426 AspAlaPheGlyIleMetAspTyrThrLeuThrGlnGlyValGlnAspAsnHisAspAsp 445
 QY 1113 ACCAACCGTATCATTAACGTTTATATGCGAACGCAACCGAGAGAGAGAAATGCTAGCTAC 1172
 Db 446 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGlnGlyGlnAsnAlaSerTyr 465
 QY 1173 CATTAGCT 1181
 Db 466 HisLeuAla 468
 RESULT 14
 US-07-854-596B-35
 : Sequence 35, Application US/07854596B
 : Patent No. 5434073
 : GENERAL INFORMATION:
 : APPLICANT: Dawson, Keith M
 : APPLICANT: Hunter, Michael G
 : APPLICANT: Czaplowski, Lloyd G
 : TITLE OF INVENTION: Proteins and nucleic acids
 : NUMBER OF SEQUENCES: 73
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Dr. John J. McDonnell
 : STREET: Ten South Wacker Drive, Suite 3000
 : CITY: Chicago
 : STATE: IL
 : COUNTRY: USA
 : ZIP: 60606
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible

: OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patent in Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/07/854,596B
 : FILING DATE: 03-JUN-1992
 : CLASSIFICATION: 435
 : ATTORNEY/AGENT INFORMATION:
 : NAME: McDonnell, John J
 : REGISTRATION NUMBER: 26,949
 : REFERENCE/DOCKET NUMBER: 92,337
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 312-715-1000
 : TELEFAX: 312-715-1234
 : TELEX: 910-221-5317
 : INFORMATION FOR SEQ ID NO: 35:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 859 amino acids
 : TYPE: amino acid
 : TOPOLOGY: linear
 : MOLECULE TYPE: protein
 : US-07-854-596B-35
 : Alignment Scores:
 : Pred. No.: 1,79e-189 Length: 859
 : Score: 1958.00 Matches: 378
 : Percent Similarity: 98.96% Conservative: 1
 : Best Local Similarity: 98.69% Mismatches: 4
 : Query Match: 72.09% Indels: 0
 : DB: 1 Gaps: 0
 : US-09-940-235-9 (1-1541) x US-07-854-596B-35 (1-859)
 :
 QY 33 ATAGCTGCTGCTGAATGGCTACTAGATCGTCTTCTGTAATTAACACCAATGCTGTGT 92
 Db 22 IleAlaIleProGlnTyrPheLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal 41
 QY 93 AGCGTGTCTGTACTGTGAGGAGCAATCAAGCAATTAATTTTGAATC 152
 Db 42 SerValAlaGlnThrValGlnGlyThrAsnGlnAspIleSerLeuLysPhePheGlnIle 61
 QY 153 GATCTAACATCAGACCTGCTCATGAGAGAAAGACAGAGCAAGGCTTAAGTCAAAATCA 212
 Db 62 AspLeuThrSerThrArgProAlaHisGlyGlyLysThrGlnGlnIleuLeuSerProLysSer 81
 QY 213 AAACCATTTCTACTGATAGTGGCGCATGTCACTATAAAGCTTGAAGAACTGACTTACTA 272
 Db 82 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGlnLysAlaAspLeuLeu 101
 QY 273 AAGGCTATTCAAGAAACATTTGATCGCTAACGTCACAGTAAAGCACTTGAAGTC 332
 Db 102 LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGlnVal 121
 QY 333 ATTGATTTTTCAGACGATGCAACCATTAATGATGCAAGCAAGGCTACTTTGCTGAC 392
 Db 122 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyValIleTyrPheAlaAsp 141
 QY 393 AAAGATGTTTGGTAACTTGGCCGACCAACCTGTCCAGAAATTTTGTAAAGGACAT 452
 Db 142 LysAspGlySerValThrLeuProThrGlnProValGlnGlnPheLeuLeuSerGlyHis 161
 QY 453 GTGCGGTTAGACCATATTAAGAAAAACAATACAAACCAAGCAATCTGTGATGTG 512
 Db 162 ValArgValArgProTyrIleGluGlyProIleGlnAsnGlnIleAlaLysSerValAspVal 181
 QY 513 GAATATACGCTAGATTTTACTCCCTTAAACCTGATGACGATTTTCAAGCAAGCTTCAAA 572
 Db 182 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 201
 QY 573 GATATACGCTATTTGAAAAACATAGCTATCGTGACCACTGACATCTCAAGAAATTA 632
 Db 202 AspThrIleValLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeu 221
 QY 633 GCTCAAGCAAAAGCAATTTTAAACAAACCAACCAAGCTATACGATTTTGAACGTGAC 692

Db 222 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp 241
QY TCCTCAATCGTCACATCATGACATGACATTTTCGTCAGCATTTTACCAATGGATCAAGAG 752
Db 242 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 261
QY TTTACTTACCGTGTAAATAATCGGAACAAGCTTATAGGATCAATAAAAAATCTGCTCG 812
Db 262 PheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsnLysLysSerGlyLeu 281
QY 813 AATGAAGAAATAACAACACATGACCTGATCTCTGAGAAATATTTACGTCCTTAAAAAGGG 872
Db 282 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 301
QY 873 GAAAGCCGTATGATCCCTTTGATCGCAGTCATCTGAAACTGTTACCATCAATAATACGTT 932
Db 302 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 321
QY 933 GATGTCGATACCAACGAATTGCTAAAGAGTGAGCAGCTCTTAACAGCTAGCGAAGCTAAC 992
Db 322 AspValAsnThrAsnGluLeuLysSerGluIleLeuThrAlaSerGluArgAsn 341
QY 993 TTAGACTTCAGAGATTATACGATCCTCGTGATAGGCTTAAACTCTCTACAACTCTC 1052
Db 342 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 361
QY 1053 GATGCTTTTGGTATTATGACATATACCTTAACCTGGAAGATGAGAGGATATACGATGAC 1112
Db 362 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 381
QY 1113 ACCAACCTATCATACACGCTTTATATGGCAGCGACCGAGGAGAGATGCTAGCTAC 1172
Db 382 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 401
QY 1173 CATTTAGCT 1181
Db 402 HisLeuAla 404

RESULT 15
US-08-488-940-18
; Sequence 18, Application US/08488940
; Patent No. 5854049
; GENERAL INFORMATION:
; APPLICANT: Reed, Guy L.
; TITLE OF INVENTION: PLASMIN-RESISTANT STREPTOKINASE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 245 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,940
; FILING DATE: 09-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 05433/009001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 1194 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-488-940-18
Alignment Scores:
Pred. No.: 2 53e-187 Length: 1194
Score: 1937.50 Matches: 390
Percent Similarity: 88.64% Conservative: 8
Best Local Similarity: 86.86% Mismatches: 45
Query Match: 71.34% Indels: 6
DB: 3 Gaps: 3
US-09-940-235-9 (1-1541) x US-08-488-940-18 (1-1194)
QY 33 ATAGCTGGTCTCGAATGGCTACTAGATCGTCCTTCTGTAATAATACACGCCAATTTGGTTGTT 92
Db 382 IleAlaGlyProGluTrpLeuLeuAspAlaProSerValAsnAsnSerGlnLeuValVal 401
QY 93 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTTTGAATC 152
Db 402 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuAlaPhePheGluIle 421
QY 153 GATCTAACATCACGACCTGCTCATGGAGGAAACACACAGCAGCAAGCTTAAAGTCCAAATCA 212
Db 422 AspLeuThrSerAlaProAlaHisGlyGlyAlaThrGluGlnGlyLeuSerProAlaSer 441
QY 213 AAACCAATTTGCTACTGATAGTGGCGGATGTCACATAAACTTGAGAAAGCTGACTTACTA 272
Db 442 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 461
QY 273 AAGCTATTCAAGAACAAATTTGATCGCTAAGCTCCACAGTAAACGACGACTCTTTGAGGTC 332
Db 462 LysAlaIleGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 481
QY 333 ATTGATTTTGCAGCGATGCAACCATTTACTGATCGAAACCGCAAGGCTTACTTTGCTGAC 392
Db 482 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 501
QY 393 AAAGATGGTTCGGTAACTTTCGCCGACCAACCTGTCCAAGAAATTTTGTAAAGCGGACAT 452
Db 502 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 521
QY 453 GTGCGCTTAGACCATATATAAGAAAAACCAATACAAACCAAGCGAAATCTGTGATGTG 512
Db 522 ValArgValArg--TyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 540
QY 513 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACCATTTTCAGACCAGGCTCATA 572
Db 541 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 560
QY 573 GATACTAAGCTATTGAAACACATAGCTATCGGTGACACCATCATCTCAAGAAATTAATA 632
Db 561 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 580
QY 633 GCTCAAGCAGCAAAAGCAATTTTAAACAAAAACCCAGGCTATACGATTTATGACCGTAC 692
Db 581 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 600
QY 693 TCCTCAATCGTCACTCATGACATGATTTTCGTCAGCATTTTACCAATGGATCAAGAG 752
Db 601 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 620
QY 753 TTTTACTTACCGTGTAAATAATCGGAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 812
Db 621 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 640
QY 813 AATGAAGAAATAACAACACATGACCTGATCTCTGAGAAATATTTACGTCCTTAAAAAGGG 872
Db 641 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 660

```
QY 873 GAAAGCCGTATGATCCCTTTGATCGCAGTCACCTTGAAACTGTTCCACCAATACGTT 932
Db |||||||
QY 661 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 680
Db |||||||
QY 933 GATGTCGATACCAACGAATTTGCTAAAGAGAGCAGCTCTTAACAGCTAGCGACGTAAC 992
Db |||||||
QY 681 AspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 700
Db |||||||
QY 993 TTAGACTTCAGAGATTATACGATCCTCGTGATAAGGCTAAAGCTACTCTACAACAATCTC 1052
Db |||||||
QY 701 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 720
Db |||||||
QY 1053 GATGCTTTTGGTATTATGACTATACCTTAACCTGGAAAAGTAGAGGATAATCACGATGAC 1112
Db |||||||
QY 721 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 740
Db |||||||
QY 1113 ACCAACCGTATCATACCGTTTATATGCGCAAGCGACCCGAGGAGAGATGCTAGCTAC 1172
Db |||||||
QY 741 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 760
Db |||||||
QY 1173 CATTTAGCTGCTGGTGGCCAGCGCAACAGATTGTACCCATAGCTGAGAAGTGTTTTGAT 1232
Db |||||||
QY 761 HisLeuAla-----TyrAspAlaAspArgTyrThrGluGluArgGluValTyrSer 778
Db |||
QY 1233 CATGCTGCTGGGACTTCCTATGCTGCGGAGAAACGTTGGGAGAGCCCTACCAAGGCTGG 1292
Db |||
QY 779 TyrLeuArgTyrThrGlyThrProIleProAspAsnProAsnAspLysAsnAsnSerGln 798
Db |||
QY 1293 ATGATGGTAGATTGT-----ACTTGCCTGGGAGAGGCGGACGCGCATCATTGCACT 1346
Db |||
QY 799 LeuValValSerValAlaGlyThrValGluGlyThrAsnGln-AspIleSerLeuLysPh 818
Db |||
QY 1347 TCTAGAAATAGATGCAACGATCAGG 1371
Db |||
QY 818 ePheGluIleAppLeuThrSerArg 826
Db |||
```

Search completed: January 28, 2006, 02:40:58
Job time : 40.9832 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 28, 2006, 01:51:16 ; Search time 100.774 Seconds
(without alignments)
12778.538 Million cell updates/sec

Title: US-09-940-235-9
Perfect score: 2716
Sequence: 1 ttgtttaacttaagaagg.....ccttcaccgatgttcgtag 1541

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 3735138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_p/US09940235/runat_27012006.144219.27635/app_query.fasta_1.7708
-DB=Published Applications AA Main -OPMT=fasta -SUFFIX=rapbm -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-MAXLEN=2000000000 -USER=US09940235 @CGN_1.1.905 @runat_27012006.144219.27635
-NCPU=6 -ICPU=3 -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DIV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA Main:

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1984	73.0	413	4	US-10-360-101-264
2	1984	73.0	413	4	Sequence 264, App
3	1963	72.3	415	5	US-09-940-235-2
4	1958	72.1	414	4	US-10-988-943-1
5	1947	71.7	414	4	US-10-300-215-252
6	1928	71.0	413	3	US-10-300-215-253
7	1794	66.1	440	5	US-09-919-703-12
8	632	23.3	259	3	US-10-474-792-658
9	632	23.3	463	4	US-09-940-235-4
10	632	23.3	463	5	US-10-144-194A-52
11	632	23.3	642	4	US-10-491-566-52
12	632	23.3	642	5	US-10-741-601-354
					Sequence 354, App
					Sequence 1066, App

13	632	23.3	657	4	US-10-741-601-359	Sequence 359, App
14	632	23.3	657	5	US-10-741-600-1072	Sequence 1072, App
15	632	23.3	984	4	US-10-741-601-356	Sequence 356, App
16	632	23.3	984	5	US-10-741-600-1069	Sequence 1069, App
17	632	23.3	2220	4	US-10-236-392-4	Sequence 4, Appli
18	632	23.3	2296	4	US-10-741-601-363	Sequence 363, App
19	632	23.3	2296	5	US-10-741-600-1075	Sequence 1075, App
20	632	23.3	2320	4	US-10-279-733-8	Sequence 8, Appli
21	632	23.3	2320	4	US-10-236-392-2	Sequence 2, Appli
22	632	23.3	2328	4	US-10-171-311-64	Sequence 64, Appl
23	632	23.3	2328	4	US-10-236-031B-70	Sequence 70, Appl
24	632	23.3	2328	4	US-10-374-979-98	Sequence 98, Appl
25	632	23.3	2328	4	US-10-182-936A-98	Sequence 98, Appl
26	632	23.3	2328	5	US-10-477-238A-677	Sequence 677, App
27	632	23.3	2328	5	US-10-680-287A-677	Sequence 677, App
28	632	23.3	2328	5	US-10-477-173-677	Sequence 677, App
29	632	23.3	2355	4	US-10-144-194A-104	Sequence 104, App
30	632	23.3	2355	4	US-10-360-101-235	Sequence 235, App
31	632	23.3	2355	4	US-10-447-161-3	Sequence 3, Appli
32	632	23.3	2355	4	US-10-734-564-94	Sequence 94, Appl
33	632	23.3	2355	4	US-10-741-601-357	Sequence 357, App
34	632	23.3	2355	4	US-10-741-601-366	Sequence 366, App
35	632	23.3	2355	5	US-10-491-566-104	Sequence 104, App
36	632	23.3	2355	5	US-10-741-600-1078	Sequence 1078, App
37	632	23.3	2355	5	US-10-741-600-1078	Sequence 1078, App
38	632	23.3	2355	5	US-10-852-335A-147	Sequence 147, App
39	632	23.3	2355	5	US-10-287-436A-436	Sequence 436, App
40	632	23.3	2355	5	US-10-287-436A-1137	Sequence 1137, App
41	632	23.3	2355	6	US-11-040-130-28	Sequence 28, Appl
42	632	23.3	2386	3	US-09-961-403-1	Sequence 1, Appli
43	632	23.3	2386	4	US-10-741-601-360	Sequence 360, App
44	632	23.3	2386	4	US-10-788-792-206	Sequence 206, App
45	632	23.3	2386	5	US-10-618-281-32	Sequence 32, Appl

ALIGNMENTS

RESULT 1
US-10-360-101-264
; Sequence 264, Application US/10360101
; Publication No. US20040009550A1
; GENERAL INFORMATION:
; APPLICANT: Moll, Gert N.
; APPLICANT: Leenhouts, Cornelis J.
; TITLE OF INVENTION: Export and modification of (poly)peptide in the lantibiotic way
; FILE REFERENCE: 2183-5673
; CURRENT APPLICATION NUMBER: US/10/360,101
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: EP 02077060.8
; PRIOR FILING DATE: 2002-05-24
; NUMBER OF SEQ ID NOS: 309
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 264
; LENGTH: 413
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: sequence of streptokinase
US-10-360-101-264

Alignment Scores:			
Pred. No.:	8.21e-177	Length:	413
Score:	1984.00	Matches:	383
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	73.05%	Indels:	0
DB:	4	Gaps:	0

US-09-940-235-9 (1-1541) x US-10-360-101-264 (1-413)

QY 33 ATAGCTGGTCTCAATGGCTACTAGATCGTCCTTCTTAATAACACCAATGGTGTGT 92
Db 1 ILEAlGlyProdiutrlpLeuLeuAaspArgProSerValAsnSerGlnLeuVal 20


```
QY 453 GTGCGGTAGACCATATAAGAAAAACCAATACAAACCAAGGAAATCTGTTGATGTG 512
|
|
|
Db 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
|
|
|
QY 513 GAATATACGTACAGTTTACTCCCTTAAACCCCTGATGACCATTTTCAGACAGCTCTCAA 572
|
|
|
Db 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
|
|
|
QY 573 GATACCTAAGCTATTTGAAAAACACTAGCTATCCGTCGACACCATCATCTCAAGAAATTA 632
|
|
|
Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 200
|
|
|
QY 633 GCTCAAGCAAAAGCAATTTTAAACAAACACCCAGGCTATACGATTTATGAACGTGAC 692
|
|
|
Db 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 220
|
|
|
QY 693 TCCTCAATCGTCACTCATGACATGACATTTCCGTACGATTTTACCAATGGATCAAGAG 752
|
|
|
Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
|
|
|
QY 753 TTTTACTTACCGTGTAAATAATCGGAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 812
|
|
|
Db 241 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 260
|
|
|
QY 813 AATGAAGAAATAAACCACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 872
|
|
|
Db 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280
|
|
|
QY 873 GAAAGCCGTATGATCCCTTTGATCCGAGTCACCTGAACTGTTCCACCATCAATAGCTT 932
|
|
|
Db 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
|
|
|
QY 933 GATGTCGATACCAACCAATTTGCTTAAAGAGTGACAGCTCTTAACAGCTACGCAACGTA 992
|
|
|
Db 301 AspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320
|
|
|
QY 993 TTAGACTTCAGAGATTATACGATCTCGTGATAAGCTAAACTACTCTACAAACAATCTC 1052
|
|
|
Db 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
|
|
|
QY 1053 GATGCTTTTGGTATTATGACTATACCTTAACCTGGAAGAGTAGAGGATATCAAGTAC 1112
|
|
|
Db 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
|
|
|
QY 1113 ACCAACCGTATCATACCTTTATATGCGCAAGCGCCGAGAGAGATGCTAGCTAC 1172
|
|
|
Db 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGluGlyGluAsnAlaSerTyr 380
|
|
|
QY 1173 CATTTAGCT 1181
|
|
|
Db 381 HisLeuAla 383
```

RESULT 3

```
US-10-988-943-1
; Sequence 1, Application US/10988943
; Publication No. US20050176085A1
; GENERAL INFORMATION:
; APPLICANT: Center for Genetic Engineering and Biotechnology
; TITLE OF INVENTION: METHOD OF SELECTIVE PEPTIDE ISOLATION FOR THE
; TITLE OF INVENTION: IDENTIFICATION AND QUANTITATIVE ANALYSIS OF PROTEINS IN
; TITLE OF INVENTION: COMPLEX MIXTURES.
; FILE REFERENCE: Proteomics CU2003-269
; CURRENT APPLICATION NUMBER: US/10/988,943
; CURRENT FILING DATE: 2004-11-15
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 415
; TYPE: PRT
; ORGANISM: Streptococcus equisimilis
US-10-988-943-1
Alignment Scores:
```

```
Pred. No.: 7,66e-175 Length: 415
Score: 1963.00 Matches: 379
Percent Similarity: 98.96% Conservative: 1
Best Local Similarity: 98.70% Mismatches: 4
Query Match: 72.28% Indels: 0
DB: 5 Gaps: 0
```

US-09-940-235-9 (1-1541) x US-10-988-943-1 (1-415)

```
QY 30 ATGATAGCTGCTGCTGCTGAATGCTACTAGATCGTCTCTCTGCTTAATAACAGCAATTTGTT 89
|
|
|
Db 1 MetIleAlaGlyProGluTyrLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
|
|
|
QY 90 GTTAGCTGCTGCTGCTACTGTTGAGCGACGAATCAAGACATTAGTCTTTAAATTTTTTGAA 149
|
|
|
Db 21 ValSerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGlu 40
|
|
|
QY 150 ATCGATCTAACATCACGACCTGCTCATCGAGGAAGACAGACGAGGCTTAAGTCCAAAA 209
|
|
|
Db 41 IleAspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLys 60
|
|
|
QY 210 TCAAAACCATTTGCTACTGATAGTGGCGCGATGTCACATAAATTGAGAAAGCTGACTTA 269
|
|
|
Db 61 SerLysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGluLysAlaAspLeu 80
|
|
|
QY 270 CTAAGGCTATTCAAGAACAAATTGATCGCTAAACGTCACAGTAACGACGACTACTTTGAG 329
|
|
|
Db 81 LeuLysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGlu 100
|
|
|
QY 330 GTCAATTGATTTGCAAGCGATGCAACCAATTACTGATGAAACGCAAGGCTACTTTGCT 389
|
|
|
Db 101 ValIleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAla 120
|
|
|
QY 390 GACAAAGATGTTGCGTAAACCTTGCCGACCCCAACCTGTCCAAAGAAATTTTGTAAAGCGGA 449
|
|
|
Db 121 AspLysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGly 140
|
|
|
QY 450 CATGTGCGCTTAGACCATATAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGAT 509
|
|
|
Db 141 HisValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAsp 160
|
|
|
QY 510 GTGGAATATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTCAGACAGGCTC 569
|
|
|
Db 161 ValGluTyrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeu 180
|
|
|
QY 570 AAGATATCTAAGCTATTGAAAAACACTAGCTATCGGTGACCACTACATCTCAAGAAATTA 629
|
|
|
Db 181 LysAspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 200
|
|
|
QY 630 CTAGCTCAAGCACAAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTATGAACGT 689
|
|
|
Db 201 LeuAlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArg 220
|
|
|
QY 690 GACTCTCTCAATCTGCTCATGACAAATGACATTTTCGTACGATTTTACCAATGATGATCAA 749
|
|
|
Db 221 AspSerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGln 240
|
|
|
QY 750 GAGTTTACTTACCGTGTATAAAATCGGGAACAAAGCTTTATAGGATCAATAAAAAATCTGGT 809
|
|
|
Db 241 GluPheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsnLysLysSerGly 260
|
|
|
QY 810 CTGAATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAA 869
|
|
|
Db 261 LeuAsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLys 280
|
|
|
QY 870 GGGGAAAGCCGTATGATCCCTTTGATCGCAGTCACCTTGAACCTGTAACCTCAACATAC 929
|
|
|
Db 281 GlyGluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyr 300
|
|
|
QY 930 GTTGATGTCATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCAACGT 989
|
|
|
Db 301 ValAspValAsnThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArg 320
```

Qy	990	AACTAGACTTCAGAGATTTTATACGATCTCGTGATAAAGCTAAACTACTCTACACAAT	1049
Db	321	AsnLeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsn	340
Qy	1050	CTCGATGCTTTTGGTATTATGGACTATACCTTAACTCGAAAGTAGAGGATATCAGAT	1109
Db	341	LeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAsp	360
Qy	1110	GACACCAACGGTATCATAAACCGTTTATATGGCGAAGCGACCCGGAAGGAGAGATGCTAGC	1169
Db	361	AspThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSer	380
Qy	1170	TACCATTTAGCT	1181
Db	381	TyrHisLeuAla	384

RESULT 4

```

US-10-300-215-252
; Sequence 252, Application US/10300215
; Publication No. US20030153043A1
; GENERAL INFORMATION:
; APPLICANT: CARR, Francis Joseph
; APPLICANT: ADAIR, Fiona Suzanne
; APPLICANT: HAMILTON, Anita Anne
; APPLICANT: CARTER, Graham
; TITLE OF INVENTION: METHOD FOR THE PRODUCTION OF
; TITLE OF INVENTION: NON-IMMUNOGENIC PROTEINS
; FILE REFERENCE: MER-104-Con.1
; CURRENT APPLICATION NUMBER: US/10/300,215
; CURRENT FILING DATE: 2002-11-20
; PRIOR APPLICATION NUMBER: US 09/438,136
; PRIOR FILING DATE: 1999-11-10
; PRIOR APPLICATION NUMBER: WO PCT/GB98/01473
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: GB 9710480.6
; PRIOR FILING DATE: 1997-05-21
; PRIOR APPLICATION NUMBER: GB 9716197.0
; PRIOR FILING DATE: 1997-07-31
; PRIOR APPLICATION NUMBER: GB 9725270.4
; PRIOR FILING DATE: 1997-11-28
; PRIOR APPLICATION NUMBER: GB 9807751.4
; PRIOR FILING DATE: 1998-04-14
; PRIOR APPLICATION NUMBER: US 60/067,235
; PRIOR FILING DATE: 1997-12-02
; NUMBER OF SEQ ID NOS: 254
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 252
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Streptococcus equisimilis
US-10-300-215-252

```

Alignment Scores:

Pred. No.:	2.25e-174	Length:	414
Argument Scores:		Matches:	378
Score:	1958.00	Conservative:	1
Percent Similarity:	98.96%	Mismatches:	4
Best Local Similarity:	98.69%	Indels:	0
Query Match:	72.09%	Gaps:	0
DB:	4		
US-09-940-235-9 (1-1541) x US-10-300-215-252 (1-414)			

33	ATAGCTGGTCTCGAATGCTACTAGATCGTCTCTTGTAATAACAGCAATGGTGTGT	92
Db	1 IlealaGlyProGluTtrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal	20
Qy	93 AGCGTTGCTGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAATTTTGGAAATC	152
Db	21 SerValalaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLyspheGluIle	40
Qy	153 GATCTAACTACGACCTGCTCTGTCGAGGAAAGACAGACAGCAGGCTTAAGTCCAAATCA	212
Db	41 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer	60

Qy	213	AAACCAATTTGCTACTGATAGTGGCGCGAGTGCACATAAACTTGTGAGAAAGCTGACTTACTA	272
Db	61	LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu	80
Qy	273	AAGCGTATTCAAGAACCAATTGATCGCTAAACGTCCACAGTACGACGACTACTCTTTGAGTC	332
Db	81	LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal	100
Qy	333	ATTGATTTTGCACGCGATGCAACCAATTACTGATCGAAACCGCAAGGCTCTACTTTGCTGCAC	392
Db	101	IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp	120
Qy	393	AAAGATGGTTCGGTAACTTTCGCCGACCACTGCTCCAGAAATTTTTCGCTAAGCGGACAT	452
Db	121	LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis	140
Qy	453	GTGCGCGTTAGACCATATAAGAAAAACCAATACAAACCAACGCAATCTGTTGATCTG	512
Db	141	ValArgValArgProTyrLysGluYspIleGlnAsnGlnAlaLysSerValAspVal	160
Qy	513	GAATPATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTCAGACCAAGTCTCAAA	572
Db	161	GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys	180
Qy	573	GATCTAAGCTATTGAAAACACTAGCTTATCGGTGACACCATCATCTCTCAGAAATTACTA	632
Db	181	AspThrLysLeuLeuYsThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu	200
Qy	633	GCTCAAGCACAAAGCAATTTTAAACAAAAAAACACCCAGGCTATACGATTTATGAACGTGAC	692
Db	201	AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp	220
Qy	693	TCCTCAATCGTCACTCATGACAATGACATTTTTCGGTAGCATTTTACCAATGGATCAGAG	752
Db	221	SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu	240
Qy	753	TTTACTTACCGTGTAAATAATCGGAAACAGCTTATAGGATCAATAAAAAATCTGGTCTG	812
Db	241	PheThrTyrHisValLysAsnArgglnGlnAlaTyrGluIleAsnLysLysSerGlyLeu	260
Qy	813	AATGAAGAAATAAACACACCTGATCTCTGAGAAATATTACGTCTTTAAAAAGGG	872
Db	261	AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly	280
Qy	873	GAAGAAGCGTATGATCCTTTTGATCGCAGTCACTTGAAACTGTTCCACCATCAATACGTT	932
Db	281	GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal	300
Qy	933	GATCTGCATACCACGAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGCTTAGCGAACGTAAAC	992
Db	301	AspValAsnThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn	320
Qy	993	TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACAAACATCTC	1052
Db	321	LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu	340
Qy	1053	GATGCTTTTGGTATTATGGACTATACCTTAACTCGAAAGCTAGAGGATATACAGATGAC	1112
Db	341	AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp	360
Qy	1113	ACCAACCGTATCATACCGTTTATATGGCGAAAGCGACCGCAAGGAGAGAATCTGAGTTCAC	1172
Db	361	ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr	380
Qy	1173	CATTAGCT 1181	
Db	381	HisLeuAla 383	

RESULT 5

US-10-300-215-253
; Sequence 253, Application US/10300215
; Publication No. US20030153043A1

GENERAL INFORMATION:
; APPLICANT: CARR, Francis Joseph
; APPLICANT: ADAIR, Fiona Suzanne
; APPLICANT: HAMILTON, Anita Anne
; APPLICANT: CARTER, Graham
; TITLE OF INVENTION: METHOD FOR THE PRODUCTION OF
; TITLE OF INVENTION: NON-IMMUNOGENIC PROTEINS
; FILE REFERENCE: MER-104-Con.1
; CURRENT APPLICATION NUMBER: US/10/300,215
; PRIOR FILING DATE: 2002-11-20
; PRIOR APPLICATION NUMBER: US 09/438,136
; PRIOR FILING DATE: 1999-11-10
; PRIOR APPLICATION NUMBER: WO PCT/GB98/01473
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: GB 9710480.6
; PRIOR FILING DATE: 1997-05-21
; PRIOR APPLICATION NUMBER: GB 9716197.0
; PRIOR FILING DATE: 1997-07-31
; PRIOR APPLICATION NUMBER: GB 9725270.4
; PRIOR FILING DATE: 1997-11-28
; PRIOR APPLICATION NUMBER: GB 9807751.4
; PRIOR FILING DATE: 1998-04-14
; PRIOR APPLICATION NUMBER: US 60/067,235
; PRIOR FILING DATE: 1997-12-02
; NUMBER OF SEQ ID NOS: 254
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 253
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Modified strep protein
US-10-300-215-253

Alignment Scores:
Pred. No.: 2,42e-173 Length: 414
Score: 1947.00 Matches: 376
Percent Similarity: 98.69% Conservative: 2
Best Local Similarity: 98.17% Mismatches: 5
Query Match: 71.69% Indels: 0
DB: Gaps: 0

US-09-940-235-9 (1-1541) x US-10-300-215-253 (1-414)

QY 33 ATAGTGTGCTCGAATGGCTACTAGATCGTCTCTGTAATAACAGCCAAATGGTGT 92
Db 1 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
QY 93 AGCGTTGCTGTACTGTCGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC 152
Db 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 40
QY 153 GATCTAACATCAGCACTGCTCATGGAGGAAAGACAGCAAGCTTAAGTCCAAATCA 212
Db 41 AspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 60
QY 213 AAACATTGCTACTGATAGTGGCGGATGTCATATAAATTGAGAAAGCTGACTACTA 272
Db 61 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu 80
QY 273 AAGCGTATTCAAGCAATTTGATCGTAAAGTCCAGTAAACGCACTACTTTGAGTTC 332
Db 81 LysAlaLysGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspThrPheGluVal 100
QY 333 ATTGATTTTGCAGCGATGCAACCACTTACTGATCGAAACGGCAAGCTTACTTTGCTGAC 392
Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValThrPheAlaAsp 120
QY 393 AAAGATGTTGGTAACTTGGCGACCCCACTGTCCTCAAGAAATTTTGTAAACGCGACAT 452
Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
QY 453 GTGCCGGTTAGACCATTATAAGAAAAACCAATCAAAAAACCAAGCAAAATCTGTTGATGTG 512

Db 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
QY 513 GAATATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCAAGTCTCAA 572
Db 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuLys 180
QY 573 GATACTAAGCTATTGAAAAACACTAGCTATCCGGTGACACCATCACATCTTCAAGAATTACTA 632
Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 200
QY 633 GCTCAGCACAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTATGAACGTGAC 692
Db 201 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp 220
QY 693 TCCTCAATCGTCACTCATGACATGACATTTTCGTACGATTTTACCAATGCGATCAAGAG 752
Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
QY 753 TTTACTTACCGTGTAAAAATCGGGAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 812
Db 241 PheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsnLysLysSerGlyLeu 260
QY 813 AATGAAGAAATAAAACACACTGACCTCATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 872
Db 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280
QY 873 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCCACCATCAATACGTT 932
Db 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysPheVal 300
QY 933 GATGTCGATACCAACGAATTTGCTAAAGTGAGCAGCTCTTAACAGCTAGCGACGTAAC 992
Db 301 AspValAsnThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320
QY 993 TTAGACTTCAGAGATTTTATACGATCTCTCGTGAATAGGCTAACTACTCTACAACAATCTC 1052
Db 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
QY 1053 GATGCTTTTGGTATTATGGACTATACCTTAACCTGGAAGAGTAGAGGATATACGATGAC 1112
Db 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
QY 1113 ACCAACCGTATCATACCGTTTATATCGGCAAGCGACCCGAGAGAGATGCTAGCTAC 1172
Db 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 380
QY 1173 CATTTAGCT 1181
Db 381 HisLeuAla 383

RESULT 6
US-09-919-703-12
; Sequence 12, Application US/09919703
; Patent No. US20020165129A1
; GENERAL INFORMATION:
; APPLICANT: Krystal, Gerald
; APPLICANT: Rabkin, Simon W.
; TITLE OF INVENTION: Peptides and Their Use to Ameliorate
; TITLE OF INVENTION: Cell Death
; FILE REFERENCE: 50216/003004
; CURRENT APPLICATION NUMBER: US/09/919,703
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 09/294,457
; PRIOR FILING DATE: 1999-04-19
; PRIOR APPLICATION NUMBER: US 08/759,599
; PRIOR FILING DATE: 1996-12-05
; PRIOR APPLICATION NUMBER: US 60/008,233
; PRIOR FILING DATE: 1995-12-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 413

db 280 GlyGluysProTyAspPProPheAspArgSerHisLeuylsLeuPheThrIleLysTyr 299

Qy 930 GTTCATGTCGATACCAACGAATTGCTTAAAAAGTCAGCAGCTCTTTAA CAGCTAGCGAACGT 989

[illegible]

DD 340 LE5ABRATAAFNEGTYUTTEMESABRYUTINIGLEINHGTYLUSVAIGTASBPASHTIASBP 359

```

Qy      1110 GACACCAACCGGATCATAAACCGGTTTATATGCGCAAGCGACCCGAGGAGAGAAATGCTAGC 116
Db      360 AspThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSer 379
Qy      1170 TACCAT 1175
Db      380 TyrHis 381

RESULT 7
US-10-474-792-658
; Sequence 658, Application US/10474792
; Publication No. US20040236072A1
; GENERAL INFORMATION:
; APPLICANT: Olmsted, Stephen
; APPLICANT: Zagursky, Robert
; APPLICANT: Nickbarg, Elliot
; APPLICANT: Winter, Lourie
; TITLE OF INVENTION: SURFACE PROTEINS OF STREPTOCOCCUS PYOGENES
; FILE REFERENCE: AM 100399
; CURRENT APPLICATION NUMBER: US/10/474,792
; CURRENT FILING DATE: 2003-10-14
; NUMBER OF SEQ ID NOS: 674
; SOFTWARE: Patent In version 3.0
; SEQ ID NO 658
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes
US-10-474-792-658

Alignment Scores:
Pred. No.:      5.5e-159      Length:      440
Score:          1794.00      Matches:     341
Percent Similarity: 93.99%    Conservative: 19
Best Local Similarity: 89.03%  Mismatches:  23
Query Match:     66.05%      Indels:      0
DB:              5           Gaps:        0

US-09-940-235-9 (1-1541) x US-10-474-792-658 (1-440)

Qy      33  ATAGCTGCTCCTGAATCGCTACTAGATCGCTCTCTCTAAATAACAGCCAAATGGTGTGTT 92
Db      27  ILeaIGlyTyrGlyTyrIleuProAspArgProIleAsnAsnSergInLeuValVal 46
Qy      93  AGCGTTCTCGGTACTGTGTGAGGGGACCAATCAAGACATAATAGTCTTAAATTTTGTGAATC 152
Db      47  SerMetaAlaGlyIleValGluGlyThrAspLysLysValPheIleAsnPhePheGluIle 66
Qy      153  GATCTTAACATCAGCACTGCTCATGAGGAGAAAGACAGACAGCAAGCGCTTAAGTCCAAATCA 212
Db      67  AspLeuThrSerGlnProAlaHisGlyGlyLysThrGlnGlnGlyLeuSerProLysSer 86
Qy      213  AAACCATTTGCTACTGATGTAGTGGCGGCATGTCCATAAACTTGAGAAAGCTGACTTACTA 272
Db      87  LysProPheAlaThrAspAsnGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu 106
Qy      273  AAGCTATTCAAGAACAAATTGATCGCTTAAACGTCCACAGTAAACGACACTACTTTGAGGTC 332
Db      107  LysAlaIleGlnLysGlnLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal 126

```



```
QY 333 ATTGATTTTCAAGCGATGCAACCATTAATGATGAAACGCAAGGCTACTTTGCTGAC 392
Db 127 IleAspPheAlaSerAspAlaThrThrAspArgAsnGlyLysValTyrPheAlaAsp 146
QY 393 AAGATGCTTCGTTACCTTCCGACCAACCTGTCACAGATTTTTCGCTAAGCGGACAT 452
Db 147 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuLysGlyHis 166
QY 453 GTCGGCTTAGACCATATAAGAAAAACCAATACAAACCAACCAAGCAATCTGTGATGTG 512
Db 167 ValArgValArgProTyrLysGluLysProValGlnAsnGlnAlaLysSerValAspVal 186
QY 513 GAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCAAGGCTCAAA 572
Db 187 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 206
QY 573 GATACCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTA 632
Db 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 226
QY 633 GCTCAAGCACAAGCATTTTAAACAAACCAACCCAGCTATACGATTTATGACGTCAC 692
Db 227 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp 246
QY 693 TCCTCAATCGTCACTCATGCAATGACATTTTCGTACGATTTTACCAATGGATCAAGAG 752
Db 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
QY 753 TTTACTACCGTGTAAAAATCGGGAACAAAGCTTATAGGATCAATAAAAAATCTGGCTG 812
Db 267 PheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsnProLysThrGlyIle 286
QY 813 AATGAGAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 872
Db 287 LysGluLysThrAsnAsnThrAspLeuValSerGluLysTyrTyrValLeuLysGlnGly 306
QY 873 GAAACCGGTATGATCCCTTTGATCGCAGTCACCTTGAAATGTTTACCACATCAATAGCT 932
Db 307 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326
QY 933 GATGTCATACCAACGAATTTGCTAAAGTAGCAGCTCTTAAACAGCTAGCGAAGCTAAC 992
Db 327 AspValAsnThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
QY 993 TTAGCTTCAGAGATTTATACGATCCCTCGGTGATAAGCTAAACTACTCTCAACAATCTC 1052
Db 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
QY 1053 GATGCTTTTGGTATTATGACTATACCTTAACCTGGAAGAGTAGAGGATATCAGCATGAC 1112
Db 367 AspAlaPheAspIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspLys 386
QY 1113 ACCAACCGTATCATACCGTTTATATGGCAAGCGACCCCAAGGAGAGAGATGCTAGCTAC 1172
Db 387 AsnAsnArgValThrValTyrMetGlyLysArgProLysGlyAlaLysGlySerTyr 406
QY 1173 CATTTAGCT 1181
Db 407 HisLeuAla 409
```

RESULT 8

```
US-09-940-235-4
; Sequence 4, Application US/09940235
; Publication No. US20030059921A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammarra
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
```

```
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/09/940,235
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-940-235-4
Alignment Scores:
Pred. No.: 4,33e-50 Length: 259
Score: 632.00 Matches: 117
Percent Similarity: 87.59% Conservative: 3
Best Local Similarity: 85.40% Mismatches: 11
Query Match: 23.27% Indels: 6
DB: 3 Gaps: 3
US-09-940-235-9 (1-1541) x US-09-940-235-4 (1-259)
QY 1143 AAGCGACCCGAGGAGAGAAATGCTAGCTACCATTTA-----GCTGGTGGTGGC 1190
Db 124 ArgArgProHis---GluThrGlyGlyTyrMetLeuGluCysValCysLeuGlyAsnGly 142
QY 1191 CAGCGGCAA---CAGATTGTACCATAGCTGAGAAAGTGTGTTTGTATCATGCTGGGACT 1247
Db 143 LysGlyGluTyrThrCysLysProIleAlaGluLysCysPheAspHisAlaAlaGlyThr 162
QY 1248 TCCTATCTGTCGGAGAAAGCTGGAGAACCCCTACCAAGCTGATGATGATGATGATGATGAT 1307
Db 163 SerTyrValValGlyGluThrTyrGluLysProTyrGlnGlyTyrMetMetValAspCys 182
QY 1308 ACTTGCCTGGGAGAAAGCAGCGGACGATCATCTTGCACCTCTAGAAATAGATGCAACGAT 1367
Db 183 ThrCysLeuGlyGluGlySerGlyArgIleThrCysThrSerArgAsnArgCysAsnAsp 202
QY 1368 CAGGACACAAGGACATCTATAGAAATGAGACACCTGGAGACAAAGAGGATAATCGAGGA 1427
Db 203 GlnAspThrArgThrSerTyrArgIleGlyAspThrTyrSerLysLysAspAsnArgGly 222
QY 1428 AACCTGCTCCATGTCATCTGCACAGGCAACGCGGAGAGAGAGTGAAGTGTGAGAGGCAC 1487
Db 223 AsnLeuLeuGlnCysIleCysThrGlyAsnGlyArgGlyGluTyrLysCysGluArgHis 242
QY 1488 ACCTCTCTGAGACACCATCGAGCGGATCTGGCCCTTCACCGATGTTCTG 1538
Db 243 ThrSerValGlnThrThrSerSerGlySerGlyProPheThrAspValArg 259
RESULT 9
US-10-144-194A-52
; Sequence 52, Application US/10144194A
; Publication No. US20030215809A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies Inc
; TITLE OF INVENTION: Regulated Breast Cancer Genes
; FILE REFERENCE: 3U 103 R1
; CURRENT APPLICATION NUMBER: US/10/144,194A
; CURRENT FILING DATE: 2002-06-12
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 52
; LENGTH: 463
; TYPE: PRT
; ORGANISM: Homo sapiens
```

1191 CAGGCGCAA -- CAGATTGTACCCATAGCTGAGAAAGTGTTTTTCATCATGCTGCTGGGACT 1247

[illegible]

Db 234 GlnAspThrArgThrSerTyrArgIleGlyAspThrTrpSerLysLysAspAsnArgGly 253

Db 234 GlnAspThrArgThrSerTyrArgIleGlyAspThrTrpSerLysLysAspAsnArgGly 253

```

; CURRENT APPLICATION NUMBER: US/10/741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 359
; LENGTH: 657
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-601-359

Alignment Scores:
Pred. No.: 5,81e-50 Length: 657
Score: 632.00 Matches: 117
Percent Similarity: 87.59% Conservatives: 3
Best Local Similarity: 85.40% Mismatches: 11
Query Match: 23.27% Indels: 6
DB: 4 Gaps: 3

US-09-940-235-9 (1-1541) x US-10-741-601-359 (1-657)

QY 1143 AAGCGACCCGAGGAGAGATGCTAGCTACCAATTA-----GCTGGTGGTGGC 1190
Db 155 ArgArgProHis---GluThrGlyGlyTyMetLeuGluCysValCysLeuGlyAsnGly 173
QY 1191 CAGCGCAA---CAGATTGTACCATAGCTGAGAGTGTGTTGATCATGCTGCTGGGACT 1247
Db 174 LysGlyGluThrCysLysProIleAlaGluLysCysPheAspHisAlaGlyThr 193
QY 1248 TCCTATGCTCGGAGAACGCTGGGAGAGCCCTACCAAGCTCGATGATGCTAGATTGT 1307
Db 194 SerTyrValValGlyGluThrTrpGluLysProTyrGlnGlyTrpMetMetValAspCys 213
QY 1308 ACTTGCCTGGGAGAGCGGACGACATCATCTGCACCTTCTAGAAATAGATGCAACGAT 1367
Db 214 ThrCysLeuGlyGluGlySerGlyArgIleThrCysThrSerArgAsnArgCysAsnAsp 233
QY 1368 CAGGACACAGGACATCTATAGAAATGGAGACACCTGGAGACAGGATATTCGAGGA 1427
Db 234 GlnAspThrArgThrSerTyrArgIleGlyAspThrTrpSerLysLysAspAsnArgGly 253
QY 1428 AACCTGCTCAGTGCACTCTGCACAGGCAACGCGGAGAGAGTGGAAAGTGTGAGAGGCAC 1487
Db 254 AsnLeuLeuGlnCysIleCysThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHis 273
QY 1488 ACCTCTGTGCAGACCATCGCGGATCTGGCCCTTCACCGATGTTGCT 1538
Db 274 ThrSerValGlnThrThrSerSerGlySerGlyProPheThrAspValArg 290

RESULT 14
US-10-741-600-1072
; Sequence 1072, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1072
; LENGTH: 657
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-600-1072

Alignment Scores:
Pred. No.: 5,81e-50 Length: 657
Score: 632.00 Matches: 117
Percent Similarity: 87.59% Conservatives: 3
Best Local Similarity: 85.40% Mismatches: 11
Query Match: 23.27% Indels: 6

```

```

; CURRENT APPLICATION NUMBER: US/10/741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 359
; LENGTH: 657
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-601-359

Alignment Scores:
Pred. No.: 5,81e-50 Length: 657
Score: 632.00 Matches: 117
Percent Similarity: 87.59% Conservatives: 3
Best Local Similarity: 85.40% Mismatches: 11
Query Match: 23.27% Indels: 6
DB: 4 Gaps: 3

US-09-940-235-9 (1-1541) x US-10-741-601-359 (1-657)

QY 1143 AAGCGACCCGAGGAGAGATGCTAGCTACCAATTA-----GCTGGTGGTGGC 1190
Db 155 ArgArgProHis---GluThrGlyGlyTyMetLeuGluCysValCysLeuGlyAsnGly 173
QY 1191 CAGCGCAA---CAGATTGTACCATAGCTGAGAGTGTGTTGATCATGCTGCTGGGACT 1247
Db 174 LysGlyGluThrCysLysProIleAlaGluLysCysPheAspHisAlaGlyThr 193
QY 1248 TCCTATGCTCGGAGAACGCTGGGAGAGCCCTACCAAGCTCGATGATGCTAGATTGT 1307
Db 194 SerTyrValValGlyGluThrTrpGluLysProTyrGlnGlyTrpMetMetValAspCys 213
QY 1308 ACTTGCCTGGGAGAGCGGACGACATCATCTGCACCTTCTAGAAATAGATGCAACGAT 1367
Db 214 ThrCysLeuGlyGluGlySerGlyArgIleThrCysThrSerArgAsnArgCysAsnAsp 233
QY 1368 CAGGACACAGGACATCTATAGAAATGGAGACACCTGGAGACAGGATATTCGAGGA 1427
Db 234 GlnAspThrArgThrSerTyrArgIleGlyAspThrTrpSerLysLysAspAsnArgGly 253
QY 1428 AACCTGCTCAGTGCACTCTGCACAGGCAACGCGGAGAGAGTGGAAAGTGTGAGAGGCAC 1487
Db 254 AsnLeuLeuGlnCysIleCysThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHis 273
QY 1488 ACCTCTGTGCAGACCATCGCGGATCTGGCCCTTCACCGATGTTGCT 1538
Db 274 ThrSerValGlnThrThrSerSerGlySerGlyProPheThrAspValArg 290

RESULT 14
US-10-741-600-1072
; Sequence 1072, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1072
; LENGTH: 657
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-600-1072

Alignment Scores:
Pred. No.: 5,81e-50 Length: 657
Score: 632.00 Matches: 117
Percent Similarity: 87.59% Conservatives: 3
Best Local Similarity: 85.40% Mismatches: 11
Query Match: 23.27% Indels: 6

```

DB:	5	Gaps:	3
US-09-940-235-9 (1-1541) x US-10-741-600-1072 (1-657)			
QY	1143	AAGCGACCCGAAGGAGAGAATGCTAGCTACCATTTA-----GCTGGTGGTGGC	1190
		:::	
Db	155	ArgArgProHis---GluThrGlyGlyTyrMetLeuGluCysValCysLeuGlyAsnGly	173
		:::	
QY	1191	CAGCGCGCAA---CAGATTGTACCATAGCTGAGAGAGTGTTTTGATCATGCTCTGGGACT	1247
		::: :::	
Db	174	LysGlyGluThrThrCysLysProIleAlaGluLysCysPheAspHisAlaIaGlyThr	193
		:::	
QY	1248	TCCTATGCTGTCGAGAGAAAGCTGGGAAAGCCCTACCAAGGCTGGAGTGGTAGATTGT	1307
		:::	
Db	194	SerTyrValValGlyGluThrTrpGluLysProTyrGlnGlyTrpMetValAspCys	213
		:::	
QY	1308	ACTTGCTCGGAGAGGACGCGGACGCATTCATTCGACTTCTAGAATAGATGCAACGAT	1367
		:::	
Db	214	ThrCysLeuGlyGluGlySerGlyArgIleThrCysThrSerArgAsnArgCysAsnAsp	233
		:::	
QY	1368	CAGGACACAGGACATCCTATAGATTGGAGACACCTGGAGCAGAAGGATATTCGGGA	1427
		:::	
Db	234	GlnAspThrArgThrSerTyrArgIleGlyAspThrTrpSerLysLysAspAsnArgGly	253
		:::	
QY	1428	AACCTGCTCAGATGTCATCTGCACAGGCCAACGGCCGAGGAGTGGAGTGTGAGAGGCAC	1487
		:::	
Db	254	AsnLeuLeuGlnCysIleCysThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHis	273
		:::	
QY	1488	ACTCTGTGAGACCAACATCAGCGGATCTGGCCCTTCACCGGATGTTGCT	1538
		:::	
Db	274	ThrSerValGlnThrThrSerSerGlySerGlyProPheThrAspValAArg	290

RESULT 15

```

US-10-741-601-356
; Sequence 356, Application US/10741601
; Publication No.:US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10/741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 356
; LENGTH: 984
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-601-356

```

Alignment Scores:

Pred. No.:	6.6e-50	Length:	984
Score:	632.00	Matches:	117
Percent Similarity:	87.59%	Conservative:	3
Best Local Similarity:	85.40%	Mismatches:	11
Query Match:	23.27%	Indels:	6
DB:	4	Gaps:	3

US-09-940-235-9 (1-1541) x US-10-741-601-356 (1-984)

1143	AAGCACC	CGGAGAGAG	AGANTGCTAG	TCACATTTA	-----GCTGTGTGGC	1190
QY	:::					
Db	155	ArgArgProHis	---GluThr	GlyGlyTy	MetLeuGluCysValCysLeuGlyAsnGly	173
QY	1191	CAGCGCAA	---CAGATTGT	TACCATAGCTGAG	AGTGTTTTGATCATGCTGCTGGGACT	1247
		::: :::				
Db	174	LysGlyGluThr	ThrCysLysPro	IleAlaGluLysCysPheAspHisAlaAlaGlyThr		193
QY	1248	TCCTATGTGT	CGGAGAGAA	CGTGGGAGACGCCT	ACCAGCGCTGGATGATGTTAGATTGT	1307
Db	194	SerTyrValVal	GlyGluThrTrp	GluLysProTyrGlnGlyTrpMetMetValAspCys		213

Qy	1308	ACTTGCTGGGAGAGGACGGACGGACGATCACCTTGCACTCTAGAAATAGATGCAAGCAT	1367
Db	214	ThrCysLeuGlyGluGlySerGlyArgIleThrCysThrSerArgAsnArgCysAsnAsp	233
Qy	1368	CAGGACACACAGGACATCCTATAGAAATGGAGACACCTGGAGCAAGAGGAGGATATCCAGGA	1427
Db	234	GlnAspThrArgThrSerTyrArgIleGlyAspThrTrpSerIlysAspAsnArgGly	253
Qy	1428	AACCTGCTCCAGTCATCTGCACAGGCAACGGCCGAGGAGAGTGGAGTCTGAGAGGCAC	1487
Db	254	AsnLeuLeuGlnCysIleCysThrGlyAsnGlyArgGlyGluTrpIysCysGluArgHis	273
Qy	1488	ACTCTGTGACAGCACCATCATGAGCGGATCTGGGCCCTTCACCGAGTTTCGT	1538
Db	274	ThrSerValGlnThrThrSerSerGlySerGlyProPheThrAspValArg	290

Search completed: January 28, 2006, 02:56:30
Job time : 113.774 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 28, 2006, 02:09:11 ; Search time 4.46194 Seconds
(without alignments)
7479.985 Million cell updates/sec

Title: US-09-940-235-9

Perfect score: 2716

Sequence: 1 ttgtttaacttaagaagg.....ccttcacgatgttcgttag 1541

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5

Ygapop 10.0, Ygapext 0.5

Fgapop 6.0, Fgapext 7.0

Delop 6.0, Delext 7.0

Searched: 75621 seqs, 10829074 residues

Total number of hits satisfying chosen parameters: 151242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame-n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_p/US09940235/runat_27012006_144220_27676/app.query.fasta_1.7708
-DB=Published Applications AA New -QFMT=fastan -SUFFIX=rapbn -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=ext -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09940235.ecgn_1_1 @runat_27012006_144220_27676
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-XGAPOP=6 -XGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications AA New:

1: /cgn2_6/prodata/2/pubppa/US08_NEW_PUB.pap:
2: /cgn2_6/prodata/2/pubppa/US06_NEW_PUB.pap:
3: /cgn2_6/prodata/2/pubppa/US07_NEW_PUB.pap:
4: /cgn2_6/prodata/2/pubppa/PCT_NEW_PUB.pap:
5: /cgn2_6/prodata/2/pubppa/US09_NEW_PUB.pap:
6: /cgn2_6/prodata/2/pubppa/US10_NEW_PUB.pap:
7: /cgn2_6/prodata/2/pubppa/US11_NEW_PUB.pap:
8: /cgn2_6/prodata/2/pubppa/US60_NEW_PUB.pap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1984	73.0	414	US-10-631-558-2	Sequence 2, Appli
2	1928	71.0	413	US-11-032-951-12	Sequence 12, Appl
3	632	23.3	259	US-10-631-558-4	Sequence 4, Appli
4	632	23.3	300	US-11-006-119-31	Sequence 31, Appl
5	632	23.3	642	US-10-995-561-631	Sequence 631, App
6	632	23.3	657	US-10-995-561-622	Sequence 622, App
7	632	23.3	984	US-10-995-561-629	Sequence 629, App
8	632	23.3	2296	US-10-995-561-633	Sequence 633, App
9	632	23.3	2355	US-10-995-561-623	Sequence 623, App
10	632	23.3	2355	US-10-995-561-627	Sequence 627, App

11	632	23.3	2384	6	US-10-821-234-1545	Sequence 1545, Ap
12	632	23.3	2386	6	US-10-995-561-626	Sequence 626, App
13	265.5	9.8	1341	6	US-10-995-561-621	Sequence 621, App
14	262.5	9.7	693	6	US-10-995-561-632	Sequence 632, App
15	262.5	9.7	1315	6	US-10-995-561-630	Sequence 630, App
16	262.5	9.7	1348	6	US-10-995-561-624	Sequence 624, App
17	247	9.1	1259	6	US-10-995-561-625	Sequence 625, App
18	208	7.7	1286	6	US-10-995-561-628	Sequence 628, App
19	189.5	7.0	847	6	US-10-995-561-634	Sequence 634, App
20	120	4.4	1992	7	US-11-013-759-3	Sequence 3, Appli
21	120	4.4	1992	7	US-11-013-759-13	Sequence 13, Appl
22	120	4.4	2047	7	US-11-013-759-4	Sequence 4, Appli
23	120	4.4	2047	7	US-11-013-759-7	Sequence 7, Appli
24	116.5	4.3	1647	7	US-11-052-554A-260	Sequence 260, App
25	115	4.2	21	7	US-11-032-951-7	Sequence 7, Appli
26	113	4.2	470	6	US-10-485-517-319	Sequence 319, App
27	113	4.2	801	7	US-11-174-150-29	Sequence 29, Appl
28	113	4.2	801	7	US-11-124-368A-292	Sequence 292, App
29	112	4.1	408	7	US-11-016-564-5	Sequence 5, Appli
30	109.5	4.0	1417	7	US-11-052-554A-8	Sequence 8, Appli
31	109.5	4.0	1448	6	US-10-485-517-212	Sequence 212, App
32	109	4.0	2340	7	US-11-052-554A-171	Sequence 171, App
33	108.5	4.0	463	6	US-10-793-626-960	Sequence 960, App
34	108.5	4.0	2314	7	US-11-013-759-11	Sequence 11, Appl
35	106	3.9	876	7	US-11-077-550-82	Sequence 82, Appl
36	106	3.9	876	7	US-11-077-550-106	Sequence 106, App
37	106	3.9	876	7	US-11-077-550-108	Sequence 108, App
38	106	3.9	2710	7	US-11-051-453-41	Sequence 41, Appl
39	105	3.9	824	6	US-10-821-234-1008	Sequence 1008, Ap
40	105	3.9	876	7	US-11-077-550-128	Sequence 128, App
41	105	3.9	877	7	US-11-077-550-126	Sequence 126, App
42	105	3.9	877	7	US-11-077-550-130	Sequence 130, App
43	105	3.9	881	7	US-11-077-550-124	Sequence 124, App
44	105	3.9	902	7	US-11-077-550-132	Sequence 132, App
45	105	3.9	912	7	US-11-077-550-116	Sequence 116, App

ALIGNMENTS

RESULT 1
US-10-631-558-2
; Sequence 2, Application US/10631558
; Publication No. US20050260598A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammar
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/10/631,558
; CURRENT FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US/09/940,235
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Streptococcus equisimilis
US-10-631-558-2
Alignment Scores:


```

; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/10/631,558
; CURRENT FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US/09/940,235
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-631-558-4

Alignment Scores:
Pred. No.: 2,76e-40 Length: 259
Score: 632.00 Matches: 117
Percent Similarity: 87.59% Conservative: 3
Best Local Similarity: 85.40% Mismatches: 11
Query Match: 23.27% Indels: 6
DB: Gaps: 3

US-09-940-235-9 (1-1541) x US-10-631-558-4 (1-259)
Qy 1143 AAGCGACCCGAGGAGAGAGATGCTAGCTACCATTTA-----GCTGGTGGTGGC 1190
Db 124 ArgArgProHis---GluThrGlyGlyTyrMetLeuGluCysValCysLeuGlyAsnGly 142
Qy 1191 CAGCGCAA---CAGATTGTACCCATAGCTGAGAGAGTGTGTTGATCATGCTGGGACT 1247
Db 143 LysGlyGluThrThrCysLysProIleAlaGluLysCysPheAspHisAlaAlaGlyThr 162
Qy 1248 TCCTATGTGTCGGAGAAACCTGGAGAACCTACCAAGGCTGATCATGCTAGATTGT 1307
Db 163 SerTyrValValGlyGluThrTrpGluLysProTyrGlnGlyTrpMetMetValAspCys 182
Qy 1308 ACTTGCCTGGAGAGAGCGAGCGCATCTTGCATCTTAGAAAATAGATGCAACGAT 1367
Db 183 ThrCysLeuGlyGluGlySerGlyArgIleThrCysThrSerArgAsnArgCysAsnAsp 202
Qy 1368 CAGGACACAGGACATCTATAGAAATTGGAGACACCTGGAGACAGAGGATAATCGAGGA 1427
Db 203 GlnAspThrArgThrSerTyrArgIleGlyAspThrTrpSerLysLysAspAsnArgGly 222
Qy 1428 AACCTGCTCCAGTCATCTGCACAGGCAACCGCGAGAGAGAGTGTGAGAGGAGC 1487
Db 223 AsnLeuLeuGlnCysIleCysThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHis 242
Qy 1488 ACCTCTGTGCAGACACATCATCGAGCGGATCTGGCCCTTCACCGCATGTTCTGT 1538
Db 243 ThrSerValGlnThrThrSerSerGlySerGlyProPheThrAspValArg 259

```

RESULT 4

```

US-11-006-119-31
; Sequence 31, Application US/11006119
; Publication No. US20050260691A1
; GENERAL INFORMATION:
; APPLICANT: Ndao, Momar
; APPLICANT: Ward, Brian
; APPLICANT: Caffrey, Rebecca
; APPLICANT: Spithill, Terry
; APPLICANT: Li, Hongshan
; APPLICANT: Podust, Vladimir
; APPLICANT: Perichon, Regis

```

RESULT 3

```

US-10-631-558-4
; Sequence 4, Application US/10631558
; Publication No. US20050260598A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammarra
; APPLICANT: Nihalani, Deepak

```

```

; APPLICANT: CIPHERGEN BIOSYSTEMS, INC.
; TITLE OF INVENTION: Serum Biomarkers for Chagas Disease
; FILE REFERENCE: 016866-012130US
; CURRENT APPLICATION NUMBER: US/11/006,119
; CURRENT FILING DATE: 2004-12-06
; PRIOR APPLICATION NUMBER: US 60/527,153
; PRIOR FILING DATE: 2003-12-05
; PRIOR APPLICATION NUMBER: US 60/565,093
; PRIOR FILING DATE: 2004-04-22
; PRIOR APPLICATION NUMBER: US 60/625,519
; PRIOR FILING DATE: 2004-11-06
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 31
; LENGTH: 300
; TYPE: PRP
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: N-terminal 28.7 fragment of Fibronectin
US-11-006-119-31

```

```

Alignment Scores:
Pred. No.: 2,81e-40 Length: 300
Score: 632.00 Matches: 117
Percent Similarity: 87.59% Conservative: 3
Best Local Similarity: 85.40% Mismatches: 11
Query Match: 23.27% Indels: 6
DB: Gaps: 3

```

US-09-940-235-9 (1-1541) x US-11-006-119-31 (1-300)

```

QY 1143 AAGCGACCCGAGGAGAGATGCTAGTACCATTTA-----GCTGGTGGTGC 1190
Db 124 ArgArgProHis---GluThrGlyGlyTrpMetLeuGluCysValCysLeuGlyAsnGly 142
QY 1191 CAGGCGCAA---CAGATTGTACCATAGTGTGAGAGTGTGTTGATCATCTGCTGGGACT 1247
Db 143 LysGlyGluTrpThrCysLysProIleAlaGluLysCysPheAspHisAlaAlaGlyThr 162
QY 1248 TCCTATGTGTCGGAGAAACGTGGAGAACGCTACCAAGGTGGATGATGTTAGATTGT 1307
Db 163 SerTyrValValGlyGluThrTrpGluLysProTyrGlnGlyTrpMetMetValAspCys 182
QY 1308 ACTTGCTGGGAGGAGCGGACGATCATCTGCACTTCTAGAAATAGATGCAACGAT 1367
Db 183 ThrCysLeuGlyGluGlySerGlyArgIleThrCysThrSerArgAsnArgCysAsnAsp 202
QY 1368 CAGGACACAGGACATCTATAGATTGGAGACACCTGGAGCAAGAGGATAATCGAGGA 1427
Db 203 GlnAspThrArgThrSerTyrArgIleGlyAspThrTrpSerLysLysAspAsnArgGly 222
QY 1428 AACCTGCTCCAGTGCATCTGCACAGGCAACCGCCGAGGAGAGTGGAGTGTGAGAGGCAC 1487
Db 223 AsnLeuLeuGlnCysIleCysThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHis 242
QY 1488 ACCTCTGTGCAGACACATCGAGCGGATCTGGCCCTTCCCGATGTTCTG 1538
Db 243 ThrSerValGlnThrThrSerSerGlySerGlyProPheThrAspValArg 259

```

RESULT 5

```

; Sequence 631, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0

```

```

; SEQ ID NO 631
; LENGTH: 642
; TYPE: PRP
; ORGANISM: Homo sapiens
US-10-995-561-631
Alignment Scores:
Pred. No.: 3.11e-40 Length: 642
Score: 632.00 Matches: 117
Percent Similarity: 87.59% Conservative: 3
Best Local Similarity: 85.40% Mismatches: 11
Query Match: 23.27% Indels: 6
DB: Gaps: 3

```

US-09-940-235-9 (1-1541) x US-10-995-561-631 (1-642)

```

QY 1143 AAGCGACCCGAGGAGAGATGCTAGTACCATTTA-----GCTGGTGGTGC 1190
Db 155 ArgArgProHis---GluThrGlyGlyTrpMetLeuGluCysValCysLeuGlyAsnGly 173
QY 1191 CAGGCGCAA---CAGATTGTACCATAGTGTGAGAGTGTGTTGATCATCTGCTGGGACT 1247
Db 174 LysGlyGluTrpThrCysLysProIleAlaGluLysCysPheAspHisAlaAlaGlyThr 193
QY 1248 TCCTATGTGTCGGAGAAACGTGGAGAACGCTACCAAGGTGGATGATGTTAGATTGT 1307
Db 194 SerTyrValValGlyGluThrTrpGluLysProTyrGlnGlyTrpMetMetValAspCys 213
QY 1308 ACTTGCTGGGAGGAGCGGACGATCATCTGCACTTCTAGAAATAGATGCAACGAT 1367
Db 214 ThrCysLeuGlyGluGlySerGlyArgIleThrCysThrSerArgAsnArgCysAsnAsp 233
QY 1368 CAGGACACAGGACATCTATAGATTGGAGACACCTGGAGCAAGAGGATAATCGAGGA 1427
Db 234 GlnAspThrArgThrSerTyrArgIleGlyAspThrTrpSerLysLysAspAsnArgGly 253
QY 1428 AACCTGCTCCAGTGCATCTGCACAGGCAACCGCCGAGGAGTGGAGTGTGAGAGGCAC 1487
Db 254 AsnLeuLeuGlnCysIleCysThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHis 273
QY 1488 ACCTCTGTGCAGACACATCGAGCGGATCTGGCCCTTCCCGATGTTCTG 1538
Db 274 ThrSerValGlnThrThrSerSerGlySerGlyProPheThrAspValArg 290

```

RESULT 6

```

; Sequence 622, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 622
; LENGTH: 657
; TYPE: PRP
; ORGANISM: Homo sapiens
US-10-995-561-622

```

Alignment Scores:

```

Pred. No.: 3.12e-40 Length: 657
Score: 632.00 Matches: 117
Percent Similarity: 87.59% Conservative: 3
Best Local Similarity: 85.40% Mismatches: 11
Query Match: 23.27% Indels: 6
DB: Gaps: 3

```

US-09-940-235-9 (1-1541) x US-10-995-561-622 (1-657)


```

US-10-995-561-623
; Sequence 623, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 623
; LENGTH: 2355
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-995-561-623

Alignment Scores:
Pred. No.: 3 7e-40 Length: 2355
Score: 632.00 Matches: 117
Percent Similarity: 87.59% Conservative: 3
Best Local Similarity: 85.40% Mismatches: 11
Query Match: 23.27% Indels: 6
DB: 6 Gaps: 3

US-09-940-235-9 (1-1541) x US-10-995-561-623 (1-2355)
QY 1143 AAGCGACCCGAGGAGAGATGCTAGCTACCATTTA-----GCTGGTGGTGGC 1190
Db 155 ArgArgProHis---GluThrGlyGlyTyrMetLeuGluCysValCysLeuGlyAsnGly 173
QY 1191 CAGGCGCAA---CAGATTGTCACCATGAGTGTGTTTGTATCATGCTGCTGGGACT 1247
Db 174 LysGlyGluThrThrCysLysProIleAlaGluLysCysPheAspHisAlaAlaGlyThr 193
QY 1248 TCCTATGTGTCGGAGAAACGTGGAGAACCCCTACCAAGCTGGATGGTAGATTGT 1307
Db 194 SerTyrValValGlyGluThrTrpGluLysProTyrGlnGlyTrpMetValAspCys 213
QY 1308 ACTTGCTGGGAGAGCGGACGCGATCCTGACCTTCTAGAAATAGATCAACGAT 1367
Db 214 ThrCysLeuGlyGluGlySerGlyArgIleThrCysThrSerArgAsnArgCysAsnAsp 233
QY 1368 CAGGACACAGGACATCCTATAGAAATTTGGAGACACCTGGAGCAAGAGGATAATCGAGGA 1427
Db 234 GlnAspThrArgThrSerTyrArgIleGlyAspThrTrpSerLysLysAspAsnArgGly 253
QY 1428 AACCTGCTCCAGTGCATCTGCACAGGCAACGCGCGAGGAGAGTGGAAGTGTGAGAGGCAC 1487
Db 254 AsnLeuLeuGlnCysIleCysThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHis 273
QY 1488 ACCTCTGTGAGACACATCGACGGATCTGGCCCTTCCACCGATGTCGT 1538
Db 274 ThrSerValGlnThrThrSerSerGlySerGlyProPheThrAspValArg 290

RESULT 10
US-10-995-561-627
; Sequence 627, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 627
; LENGTH: 2355
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-995-561-627

Alignment Scores:
Pred. No.: 3 7e-40 Length: 2384
Score: 632.00 Matches: 117
Percent Similarity: 87.59% Conservative: 3
Best Local Similarity: 85.40% Mismatches: 11
Query Match: 23.27% Indels: 6
DB: 6 Gaps: 3

```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-995-561-627

Alignment Scores:
Pred. No.: 3 7e-40 Length: 2355
Score: 632.00 Matches: 117
Percent Similarity: 87.59% Conservative: 3
Best Local Similarity: 85.40% Mismatches: 11
Query Match: 23.27% Indels: 6
DB: 6 Gaps: 3

US-09-940-235-9 (1-1541) x US-10-995-561-627 (1-2355)
QY 1143 AAGCGACCCGAGGAGAGATGCTAGCTACCATTTA-----GCTGGTGGTGGC 1190
Db 155 ArgArgProHis---GluThrGlyGlyTyrMetLeuGluCysValCysLeuGlyAsnGly 173
QY 1191 CAGGCGCAA---CAGATTGTCACCATGAGTGTGTTTGTATCATGCTGCTGGGACT 1247
Db 174 LysGlyGluThrThrCysLysProIleAlaGluLysCysPheAspHisAlaAlaGlyThr 193
QY 1248 TCCTATGTGTCGGAGAAACGTGGAGAACCCCTACCAAGCTGGATGGTAGATTGT 1307
Db 194 SerTyrValValGlyGluThrTrpGluLysProTyrGlnGlyTrpMetValAspCys 213
QY 1308 ACTTGCTGGGAGAGCGGACGCGATCCTGACCTTCTAGAAATAGATCAACGAT 1367
Db 214 ThrCysLeuGlyGluGlySerGlyArgIleThrCysThrSerArgAsnArgCysAsnAsp 233
QY 1368 CAGGACACAGGACATCCTATAGAAATTTGGAGACACCTGGAGCAAGAGGATAATCGAGGA 1427
Db 234 GlnAspThrArgThrSerTyrArgIleGlyAspThrTrpSerLysLysAspAsnArgGly 253
QY 1428 AACCTGCTCCAGTGCATCTGCACAGGCAACGCGCGAGGAGAGTGGAAGTGTGAGAGGCAC 1487
Db 254 AsnLeuLeuGlnCysIleCysThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHis 273
QY 1488 ACCTCTGTGAGACACATCGACGGATCTGGCCCTTCCACCGATGTCGT 1538
Db 274 ThrSerValGlnThrThrSerSerGlySerGlyProPheThrAspValArg 290

RESULT 11
US-10-821-234-1545
; Sequence 1545, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_SEQ_genes Version 1.0
; SEQ ID NO 1545
; LENGTH: 2384
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-821-234-1545

Alignment Scores:
Pred. No.: 3 7e-40 Length: 2384
Score: 632.00 Matches: 117
Percent Similarity: 87.59% Conservative: 3
Best Local Similarity: 85.40% Mismatches: 11
Query Match: 23.27% Indels: 6
DB: 6 Gaps: 3

```

US-09-940-235-9 (1-1541) x US-10-821-234-1545 (1-2384)

```

QY 1143 AAGCGACCCGAGGAGAGATGCTAGCTACCATTTA-----GCTGGTGGTGC 1190
Db ::::||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
184 ArgArgProHis---GluThrGlyGlyTyrMetLeuGluCysValCysLeuGlyAsnGly 202
QY 1191 CAGCGCCAA---CAGATTGTACCATAGCTGAGAGAGTGTGATCATGCTGGGACT 1247
Db :::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
203 LysGlyGluTrpThrCysLysProIleAlaGluLysCysPheAspHisAlaAlaGlyThr 222
QY 1248 TCCTATGTCGTCGAGAAAGCTGGAGAGCCCTACCAAGCTCGATGATGATGATGT 1307
Db :::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
223 SerTyrValValGlyGluThrTrpGluLysProTyrGlnGlyTrpMetMetValAspCys 242
QY 1308 ACTTGCTGGGAGAGCGACGACGATCATTGCACTTCTACAAATAGATGCAACGAT 1367
Db :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
243 ThrCysLeuGlyGluGlySerGlyArgIleThrCysThrSerArgAsnArgCysAsnAsp 262
QY 1368 CAGGACACAGGATCCTATAGAAATGGAGACACCTGGAGCAAGAGGATAATCGAGGA 1427
Db :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
263 GlnAspThrArgThrSerTyrArgIleGlyAspThrTrpSerLysAspAsnArgGly 282
QY 1428 AACCTGCTCAGTGCATCTCCACAGGCAACGCGGAGAGTGAAGTGTGAGAGGCAC 1487
Db :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
283 AsnLeuLeuGlnCysIleCysThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHis 302
QY 1488 ACTCTGTGTCAGACCATCGAGCGGATCTGGCCCTTCACCGATGTTTCT 1538
Db :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
303 ThrSerValGlnThrThrSerSerGlySerGlyProPheThrAspValArg 319

```

RESULT 12

```

US-10-995-561-626
; Sequence 626, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; NUMBER OF SEQ ID NOS: 2004-11-24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 626
; LENGTH: 2386
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-626

```

```

Alignment Scores:
Pred. No.: 3,7e-40 Length: 2386
Score: 632.00 Matches: 117
Percent Similarity: 87.59% Conservative: 3
Best Local Similarity: 85.40% Mismatches: 11
Query Match: 23.27% Indels: 6
DB: Gaps: 3

```

US-09-940-235-9 (1-1541) x US-10-995-561-626 (1-2386)

```

QY 1143 AAGCGACCCGAGGAGAGATGCTAGCTACCATTTA-----GCTGGTGGTGC 1190
Db ::::||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
155 ArgArgProHis---GluThrGlyGlyTyrMetLeuGluCysValCysLeuGlyAsnGly 173
QY 1191 CAGCGCCAA---CAGATTGTACCATAGCTGAGAGTGTGATCATGCTGGGACT 1247
Db :::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
174 LysGlyGluTrpThrCysLysProIleAlaGluLysCysPheAspHisAlaAlaGlyThr 193
QY 1248 TCCTATGTCGTCGAGAAAGCTGGAGAGCCCTACCAAGCTCGATGATGATGATGT 1307
Db :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
194 SerTyrValValGlyGluThrTrpGluLysProTyrGlnGlyTrpMetMetValAspCys 213
QY 1308 ACTTGCTGGGAGAGCGACGACGATCATTGCACTTCTACAAATAGATGCAACGAT 1367

```

```

Db 214 ThrCysLeuGlyGluGlySerGlyArgIleThrCysThrSerArgAsnArgCysAsnAsp 233
QY 1368 CAGGACACAGGATCCTATAGAAATGGAGACACCTGGAGCAAGAGGATAATCGAGGA 1427
Db :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
234 GlnAspThrArgThrSerTyrArgIleGlyAspThrTrpSerLysAspAsnArgGly 253
QY 1428 AACCTGCTCAGTGCATCTCCACAGGCAACGCGGAGAGTGAAGTGTGAGAGGCAC 1487
Db :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
254 AsnLeuLeuGlnCysIleCysThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHis 273
QY 1488 ACCTCTGTGTCAGACCATCGAGCGGATCTGGCCCTTCACCGATGTTTCT 1538
Db :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
274 ThrSerValGlnThrThrSerSerGlySerGlyProPheThrAspValArg 290

```

RESULT 13

```

US-10-995-561-621
; Sequence 621, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; NUMBER OF SEQ ID NOS: 2004-11-24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 621
; LENGTH: 1341
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-621

```

```

Alignment Scores:
Pred. No.: 8e-13 Length: 1341
Score: 265.50 Matches: 118
Percent Similarity: 33.90% Conservative: 62
Best Local Similarity: 22.22% Mismatches: 160
Query Match: 9.78% Indels: 191
DB: Gaps: 23

```

US-09-940-235-9 (1-1541) x US-10-995-561-621 (1-1341)

```

QY 72 AATAACACCAATTTGGTTGTAGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATT 131
Db :::: ||| :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
848 AspSerSerSerValValValSer-----GlyLeuMetValAlaThrLysTyrGluVal 865
QY 132 AGTCTTAATTTTTGAAATCGATCTAACATCAGACCTGCTCATGCGGAAACACAGAG 191
Db :||| ||| :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
866 SerValTyrAlaLeuLysAspThrLeuThrSerArgProAlaGlnGlyValValThr 885
QY 192 ---CAAGGCTTAAGTCCAAAATCAAAACCAATTTGCTACTGATAGTGGCGGATGTCACAT 248
Db :::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
886 LeuGluAsnValSerProArgAlaValThrAsp----- 899
QY 249 AAACCTTGAGAAAGCTGACTTACTAAAGCTATTCAAGAACAAATTCGTCGTAACGTCAC 308
Db :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
899 ----- 899
QY 309 AGTAACGACGACTACTTTGAGGTCAATTTTTCGAGCGGATGCAACCAATTCATGATCGA 368
Db :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
900 -----AlaThrGluThrThrIle----- 907
QY 369 AACGGCAAGGTCTACTTTGTCGACAAAGATGGTTGGTAACCTTGGCCGACCAACCTGTC 428
Db :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
908 -----SerTyrArgThrLysThrGluThrIle 916
QY 429 CAAGAAATTTTGTAAAGCGGACATGTGCGCGTTAGACCATATAAGAAACCAATACAA 488
Db :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
917 ThrGlyPheGlnValAspAla-----ValProAlaAsnGlyGlnThrProIleGln 933

```



```
Qy 609 ACCATCATCTCAAGAAATTACTAGCTCAAGCA----- 641
Db 874 AsnAlaArgSerProValValIleAspAlaThrAlaIleAspAlaProSerAsn 893
Qy 642 -----CAAAGCATTTTAAACAAACCAACCCCA----- 668
Db 894 LeuArgPheLeuAlaThrThrProAsnSerLeuLeuValSerTrpGlnProProArgAla 913
Qy 669 -----GGCTATACGATT--TATGAACGTGACTCCTCAATCGTCACTCATGACAAAT 716
Db 914 ArgIleThrGlyTyrIleIleLysTyrGluLysProGlySerProPro----- 929
Qy 717 GACATTTTCCGTACCAATTTTACCAATGATCAA----- 749
Db 930 -----ArgGluValValProArgProArgProGlyValThrGluAlaThrIleThr 946
Qy 750 -----GAGTTTACTTACCGTGTAAAAATCGGGAACAAGCTTATAGG 791
Db 947 GlyLeuGluProGlyThrGluTyrThrIleTyrVal-----IleAlaLeuLys 962
Qy 792 ATCAATAAANAATCT-----GGTCTGAATGAAGAAATAAACAACACTGACCTG 839
Db 963 AsnAsnGlnLysSerGluProLeuIleGlyArgLysThrValGlnLysThrProPhe 982
Qy 840 ATCTCTGAGAAATATACGCTCTTAAAAAAGG----- 872
Db 983 ValThrHisProGlyTyrAspThrGlyAsnGlyIleGlnLeuProGlyThrSerGlyGln 1002
Qy 872 ----- 872
Db 1003 GlnProSerValGlyGlnGlnMetIlePheGluGluHisGlyPheArgArgThrThrPro 1022
Qy 873 -----GAAAAGCGTATGATCCC---TTTGATCCG 899
Db 1023 ProThrThrAlaThrProIleArgHisArgProArgProTyrProProAsnValGlyGln 1042
Qy 900 AGTCACCTTGAAACTGTTCAACCATCAATACGTTGATGTC---GATACCAACGAATTGCTA 956
Db 1043 GluAlaLeuSerGlnThrThrIleSerTrpAlaProPheGlnAspThrSerGluTyrIle 1062
Qy 957 AAAAGTGAAGCTCTTAACAGCTAGCGAACGTAACCTTAGACTTCAGAGATTTATACGAT 1016
Db 1063 IleSerCysHisProValGlyThrAspGluGluProLeuGlnPheArg-----Val 1079
Qy 1017 CCTCGTGATAGGCTAACTACTCTACAACAATCTCGATGCTTTGGTATTATGACTAT 1076
Db 1080 ProGlyThrSerThrSerAla----- 1086
Qy 1077 ACCTTAACCTGGA-----AAAGTA 1094
Db 1087 ThrLeuThrGlyLeuThrArgGlyAlaThrTyrAsnIleIleValGluAlaLeuLysAsp 1106
Qy 1095 GAGGATAATCAGATGACACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAA 1154
Db 1107 GlnGlnArgHisLeuLysValArgGluGluValValThrValGlyAsnSerValAsn---Glu 1125
Qy 1155 GGAGAGAAATGCTAGTACCATTTAGCTGTGTGTGGCCAGCGCAACAGATTGTACCCATA 1214
Db 1126 GlyLeuAsnGln-----ProThr 1131
Qy 1215 GCTGAGAAGTGTGTTGATCATCTGCTGGACTTCCTATGTGTGGAGAAACGTGGAG 1274
Db 1132 AspAspSerCysPheAspProTyrThrValSerHisTyrAlaValGlyAspGluTrpGlu 1151
Qy 1275 AAGCCCTACCAAGGCTGGATGATGTAGATTGTACTTGTGGGAGAGCGACGCGACGC 1334
Db 1152 ArgMetSerGluSerGlyPheLysLeuLeuCysGlnCysLeuGlyPheGlySerGlyHis 1171
Qy 1335 ATCACTTCGACTCTTAGAAATAGATGCAACGATCAGGACACAAGGACATCCTATAGAAAT 1394
Db 1172 PheArgCysAspSerSerArgTrpCysHisAspAsnGly-----ValAsnTyrLysIle 1189
Qy 1395 GGAGACACTGGAGCAGAGAGATTAATCGAGGAAACCTGCTCCAGTGCATCTGCACAGGC 1454
```

```
Db 1190 GlyGluLysTrpAspArgGlnGlnGlyGluAsnGlyGlnMetMetSerCysThrCysLeuGly 1209
Qy 1455 AACGCCCGAGGAGAGTGGAAAGTGTGAGAGGCAC 1487
Db 1210 AsnGlyLysGlyGluPheLysCysAspProHis 1220
```

Search completed: January 28, 2006, 02:57:30
Job time : 25.4619 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 28, 2006, 01:49:35 ; Search time 23.6156 Seconds
(without alignments)
12556.953 Million cell updates/sec

Title: US-09-940-235-9
Perfect score: 2716
Sequence: 1 ttgtttaacttaagaagg.....cattaccgatgttcgtag 1541

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p.model -DEV=xl
-Q=/cgn2_1/USPTO_spool_p/US09940235/runat_27012006_144218_27578/app_query.fasta_1.7708
-DB=PIR -QFMT=fastan -SUFFIX=rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09940235 @CGN_1_1_185 @runat_27012006_144218_27578 -NCPU=6 -ICPU=3
-NO_WMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR_80.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1984	73.0	440	1 A22801	streptokinase prc
2	1951.5	71.9	415	1 B2S0	streptokinase (EC
3	1942	71.5	440	2 S02723	streptokinase G pr
4	1779	65.5	440	2 S02724	streptokinase A pr
5	1673	61.6	414	2 JU0292	streptokinase - St
6	1670	61.5	440	2 S04168	streptokinase A pr
7	667	24.6	197	2 S53334	streptokinase - St
8	632	23.3	2386	1 FNUH	fibronectin precu
9	620	22.8	128	2 S77671	streptokinase A (E
10	610	22.5	2265	1 FNB0	fibronectin - bovi
11	585.5	21.6	2477	2 S14428	fibronectin precu
12	569	20.9	128	2 S77680	streptokinase A (E
13	557	20.5	128	2 S77688	streptokinase A (E
14	552	20.3	128	2 S77679	streptokinase A (E

15	544	20.0	2481	2 A43908	fibronectin - Afri
16	423	15.6	128	2 S77676	streptokinase A (E
17	419	15.4	128	2 S77673	streptokinase A (E
18	419	15.4	128	2 S77687	streptokinase A (E
19	417	15.4	128	2 S77682	streptokinase A (E
20	416	15.3	128	2 S77678	streptokinase A (E
21	408	15.0	128	2 S77677	streptokinase A (E
22	407	15.0	128	2 S77685	streptokinase A (E
23	401	14.8	128	2 S77683	streptokinase A (E
24	398	14.7	128	2 S77686	streptokinase A (E
25	397	14.6	128	2 S77684	streptokinase A (E
26	396	14.6	128	2 S77681	streptokinase A (E
27	390	14.4	128	2 S77674	streptokinase A (E
28	386	14.2	128	2 S77675	streptokinase A (E
29	221	8.1	1020	2 A29355	fibronectin - chic
30	182	6.7	190	2 151279	fibronectin - east
31	146.5	5.4	2231	2 D71870	hypothetical prote
32	130	4.8	1745	2 A46431	tight junction-ass
33	130	4.8	2334	2 S32920	cell wall-associat
34	129.5	4.8	940	2 AD1374	internalin protein
35	126	4.6	4688	2 F82885	hypothetical prote
36	125.5	4.6	940	2 AB1744	internalin protein
37	125	4.6	1220	1 DUBEC3	DNA-directed DNA p
38	124.5	4.6	627	2 F84194	Htr14 transducer
39	124.5	4.6	1315	2 T28679	fibrinogen-binding
40	123	4.5	1019	2 T50251	hypothetical coile
41	122.5	4.5	1435	2 S54697	DNA polymerase III
42	122.5	4.5	1435	2 C90596	hypothetical prote
43	121	4.5	785	2 D71653	cell surface antig
44	121	4.5	1237	2 AC1583	internalin protein
45	121	4.5	1385	2 D89824	hypothetical prote

ALIGNMENTS

RESULT 1

A22801
streptokinase precursor - Streptococcus "equisimilis"
C;Species: Streptococcus "equisimilis"
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: A22801
R;Malke, H.; Roe, B.; Ferretti, J.J.
Gene 34, 357-362, 1985
A;Title: Nucleotide sequence of the streptokinase gene from Streptococcus equisimilis H
A;Reference number: A22801; MUID:85232082; PMID:2989113
A;Accession: A22801
A;Molecule type: DNA
A;Residues: 1-440 <MAL>
A;Cross-references: UNIPROT:P00779; UNIPARC:UPI000002BE73; GB:X72832; NID:g407876; PIDN
A;Experimental source: strain H46A
C;Genetics:
A;Gene: skc
C;Superfamily: streptokinase

Alignment Scores:
Pred. No.: 2.53e-127 Length: 440
Score: 1984.00 Matches: 383
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 73.05% Indels: 0
DB: 1 Gaps: 0

US-09-940-235-9 (1-1541) x A22801 (1-440)

QY	33	ATAGCTGCTCCTGAATGGCTACTAGATCGTCCTTCTGTAATAACAGCAATTCGTTGTT	92
DB	27	IlEaLaGlyProGluTrpLeuLeuAspArgProSerValAsnSerGlnLeuVal	46
QY	93	AGCGTTGCTGGTACTGTGTGAGGGGACGAATCAAGACATTAGTCTTAATTTTTTGAATC	152
DB	47	SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLeysPhePheGluIle	66
QY	153	GATCTTAATCATCAGCACCTGCTCATGTGAGGAGAAAGACAGACAGAGCTTAAGTCCAAATCA	212

|||||
67 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 86
|||||
213 AAACCATTTGCTACTGATAGTGGCGCATGTCCACATAAAGTTCAGAAAGCTGCTACTTA 272
|||||
87 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLysLeuGluLysAlaAspLeuLeu 106
|||||
273 AAGGCTATTCAAGAACAAATTGATCGCTAACGCTCCACAGTAAGACGACACTACTTTGAGGTC 332
|||||
107 LysAlaIleGlnGlnLeuIleAlaenValHisSerAsnAspAspTyrPheGluVal 126
|||||
333 ATTGATTTTGCAGCGATGCAACCAATTACTGATCGAAAGCGCAAGTCTACTTTGCTGAC 392
|||||
127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146
|||||
393 AAAGATGCTTCGGTAACTTCGCGACCAACCTGTCCAAGATTTTGGCTTAAGCGACAT 452
|||||
147 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 166
|||||
453 GTGCGCGTTAGACCATATAAGAAAAACCAATACAAAAACCAAGCGAAATCTGTTGATGTG 512
|||||
167 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 186
|||||
513 GAATATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGTCTCTCAA 572
|||||
187 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 206
|||||
573 GATACTAAGCTATTGAAACACTAGCTATCGTGTGACCATCACATCTCAAGATTTACTA 632
|||||
207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 226
|||||
633 GCTCAAGCACAAAGCATTTTAAACAAAAACCAACCCAGGCTATACAGATTTATCAACGTGAC 692
|||||
227 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 246
|||||
693 TCCTCAATCGTCACATGACATGACATTTTCGTTACGATTTTACCATGGATCAAGAG 752
|||||
247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
|||||
753 TTTACTTACCGTGTAAAAATCGGGAACAGCTTATAGATCAATAAAAAATCTGCTGTG 812
|||||
267 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 286
|||||
813 AATGAGAAATAAACAACACTGATCTCTGAGAAATATTAGTCCCTTAAAAAGGG 872
|||||
287 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 306
|||||
873 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACATTGAAACTCTTCAACATCAAAATACGTT 932
|||||
307 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326
|||||
933 GATGCGNATCAACGAATTCATAAAGATGAGCAGCTCTTAACAGCTAGCGAACGTAAC 992
|||||
327 AspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
|||||
993 TTAGACTTCAGATTTATACGATCTCGTGNATAGGCTTAACTACTCTACACATCTC 1052
|||||
347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
|||||
1053 GATGCTTTTGTGATTTATCGCATATACCTTAACTCGAAAAAGTAGAGGATATCAACGATGAC 1112
|||||
367 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAsnHisAspAsp 386
|||||
1113 ACCAACCGTATCATAAACCGTTTATATGGCGAAGCGACCCGGAAGGAGAGAAATCTAGCTAC 1172
|||||
387 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 406
|||||
1173 CATTTAGCT 1181
|||||
407 HisLeuAla 409
|||||

RESULT 2
BZSO

streptokinase (EC 3.4.-.-) - Streptococcus sp.
C;Species: Streptococcus sp.
C;Date: 05-Apr-1983 #sequence_revision 05-Apr-1983 #text_change 09-Jul-2004
C;Accession: A00967
R;Jackson, K.W.; Tang, J.
Biochemistry 21, 6620-6625, 1982
A;Title: Complete amino acid sequence of streptokinase and its homology with serine protease
A;Reference number: A00967; MUID:83127125; PMID:6760891
A;Accession: A00967
A;Molecule type: protein
A;Residues: 1-415 <JAC>
A;Cross-references: UNIPROT:P00779; UNIPARC:UPI00000172BF5
A;Note: 169-Asp and 181-Asp were also found
C;Superfamily: streptokinase
C;Keywords: hydrolase

Alignment Scores:
Pred. No.: 4,09e-125 Length: 415
Score: 1951.50 Matches: 380
Percent Similarity: 99.22% Conservative: 1
Best Local Similarity: 98.96% Mismatches: 2
Query Match: 71.85% Indels: 1
DB: Gaps: 1
US-09-940-235-9 (1-1541) x BZSO (1-415)

QY 33 ATAGCTGGTCTCGAATGGCTACTAGATCGTCTTCTGTAATAATACAGCCAAATGGTTGTT 92
|||||
Db 1 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
|||||
QY 93 AGCGTTCGTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAATC 152
|||||
Db 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspLysSerLeuLysPhePheGluIle 40
|||||
QY 153 GATCTAACATCAAGACCTGCTCATGGAGGAGAAACAGACAGCAAGCTTAAAGTCCAAATCA 212
|||||
Db 41 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 60
|||||
QY 213 AAACATTTGCTACTGATGATGCGCGATGTACATAAACTTGAGAAAGCTGACTTACTA 272
|||||
Db 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLysLeuGluLysAlaAspLeu 80
|||||
QY 273 AAGGCTATTCAAGAACAAATTGATCGTCAAGTCCAGTAACAGTAACAGCTACTTGGTGC 332
|||||
Db 81 LysAlaIleGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100
|||||
QY 333 ATTGATTTTGCAGCGATGCAACCAATTACTGATCGAAACCGCAAGGTCTACTTTGCTGAC 392
|||||
Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
|||||
QY 393 AAAGATGCTTCGGTAACTTCGCGACCCCAACCTGTCCAGAAATTTTTCGTAAGCGGACAT 452
|||||
Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
|||||
QY 453 GTGCGCGTTAGACCATATAAGAAAAACCAATACAAAAACCAAGCGAAATCTGTTGATGTG 512
|||||
Db 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
|||||
QY 513 GAATATACGTGATCAGATTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGTCTCAAA 572
|||||
Db 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
|||||
QY 573 GATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTTACTA 632
|||||
Db 181 LeuThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 200
|||||
QY 633 GCTCAAGCACAAAGCAATTTTAAACAAAAACCAAGCGCTATAGGATTTATGACGCTGAC 692
|||||
Db 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 220
|||||
QY 693 TCCTCAATCGTCACTCATGACATGACATTTTCGTTACGATTTTACCATGGATCAAGAG 752
|||||


```
Db      221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
Qy      753 TTTACTACCGTGTAAAAATCGGGAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 812
Db      241 PheThrTyArgValLysAsnArgGluGlnAlaTyArgIleAsnLysLysSerGlyLeu 260
Qy      813 AATGAAGAATAAACAACACTGACCTGATCTCTGAGAAATAT---TACGTCCTTAAAAA 869
Db      261 AsnGluGluIleAsnAsnThrAspLeuIleSerLeuGluTyLysTyValLeuLysLys 280
Qy      870 GGGGAAAACCGGTATCATCCCTTTGATCGCAGTCACTTGAACATGTTCCACCAATATAC 929
Db      281 GlyGluLysProTyArgProPheAspArgSerHisLeuLysLeuPheThrIleLysTy 300
Qy      930 GTTGATGTCGATACCAACGAATTTGCTAAAAGTGAGCAGCTCTTAAACAGCTAGCGAACGT 989
Db      301 ValAspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArg 320
Qy      990 AACTTAGACTTCAGAGATTATACGATCTCTCGTGATAAGGCTAAACTTACTCTACAACAAT 1049
Db      321 AsnLeuAspPheArgAspLeuTyArgProArgAspLysAlaLysLeuLeuTyAsnAsn 340
Qy      1050 CTCGATGCTTTGGTATTATGGAATACCTTAACTGCGAAGTACGAGGATATATCAGAT 1109
Db      341 LeuAspAlaPheGlyIleMetAspTyThrLeuThrGlyLysValGluAspAsnHisAsp 360
Qy      1110 GACACCAACCGGTATACACCGTTTATATGGGCAAGCGACCCGAAGGAGAGATGCTAGC 1169
Db      361 AspThrAsnArgIleIleThrValTyMetGlyLysArgProGluGlyGluAsnAlaSer 380
Qy      1170 TACCATTAGCT 1181
Db      381 TyrHisLeuAla 384

RESULT 3
S02723
streptokinase G precursor - Streptococcus sp. (strain 19908)
C:Species: Streptococcus sp.
C:Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 09-Jul-2004
C:Accession: S02723
R:Walter, F.; Siegel, M.; Maik, H.
Nucleic Acids Res. 17, 1262, 1989
A:Title: Nucleotide sequence of the streptokinase gene from a group-G Streptococcus.
A:Reference number: S02723; MUID:89160265; PMID:2922269
A:Accession: S02723
A:Molecule type: DNA
A:Residues: 1-440 <WAL>
A:Cross-references: UNIPROT:P10519; UNIPARC:UPI000013614B; EMBL:X13400; NID:g47095; PIDN
C:Genetics:
C:Superfamily: streptokinase
F:1-26/Domain: signal sequence #status predicted <SIG>
F:27-440/Product: streptokinase #status predicted <MAT>

Alignment Scores:
Pred. No.: 1-82e-124 Length: 440
Score: 1942.00 Matches: 375
Percent Similarity: 98.43% Conservative: 2
Best Local Similarity: 97.91% Mismatches: 6
Query Match: 71.50% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-9 (1-1541) x S02723 (1-440)

Qy      33 ATAGCTGGTCTGAATGGCTACTAGATCGTCTCTTCTGTAATAAACAAGCAATTTGGTTGT 92
Db      27 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal 46
Qy      93 AGCGTGTCTGGTACTGTTGAGGGGACGAATCAAGACATTAAGATCTTAAATTTTGAATC 152
Db      47 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 66
Qy      153 GATCTAACATCAGCACCTGCTCATCGAGGAGAAAGACAGACGAGCTTAAGTCCAAATCA 212
```

RESULT 4
S02724

```
Db      67 AspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 86
Qy      213 AAACCATTTGCTACTGATAGTGGCGGATGTCACATAAACTTGAGAAAGCTGACTTACTA 272
Db      87 LysLeuPheAlaThrAspSerGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu 106
Qy      273 AAGCTATTTCAGGAACAATTCATCGCTAAACCTCCACAGTACAGCAGCTACTTTGAGTGC 332
Db      107 LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspTyPheGluVal 126
Qy      333 ATTGATTTTGCACGCGATGCAACCATTTACTGATCGAAACGCAAGGTCTACTTTCTCTGAC 392
Db      127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyPheAlaAsp 146
Qy      393 AAGATGGTTCGGTAAACCTTGCAGCCCACTCTCCAGAAATTTTTCGTAAGCGGACAT 452
Db      147 LysAspGlySerValThrLeuProIleGlnProValGlnGluPheLeuLeuLysGlyHis 166
Qy      453 GTGCGCTTAGACCATATAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 512
Db      167 ValArgValArgProTyLysGluLysProValGlnAsnGlnAlaLysSerValAspVal 186
Qy      513 GAATATATCTGTACAGTTTACTCCCTTAAACCTCGATGACGATTTTCAGACCAAGTCTCAAA 572
Db      187 GluTyThrValGlnPheThrProLeuAsnProAspAspPheArgProAlaLeuLys 206
Qy      573 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTACTA 632
Db      207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 226
Qy      633 GCTCAAGCACAAAGCATTTTAAACAAAACCAACCCAGCAGCTATACGATTATGACGTCGAC 692
Db      227 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyThrIleTyGluArgAsp 246
Qy      693 TCCTCAATCGTCACTCATGCAATGACATTTTCCGTACGATTTTCCAAATGGATCAAGAG 752
Db      247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
Qy      753 TTTACTACCGTGTAAAAATCGGGAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 812
Db      267 PheThrTyHisValLysAsnArgGluGlnAlaTyArgIleAsnLysLysSerGlyLeu 286
Qy      813 AATCGAATAAACAACACTGACCTGATCTCTGAGAAATATTAACGTCCTTAAAAAGGG 872
Db      287 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyThrValLeuLysLysGly 306
Qy      873 GAAAAGCCGTATCATCCCTTTGATCGCAGTCACCTTGAAACTGTTCCACCATCAAAATACGTT 932
Db      307 GluLysProTyArgProPheAspArgSerHisLeuLysLeuPheThrIleLysTyVal 326
Qy      933 GATGTCGATACCAACGAATTTGCTAAAAGTGAGCAGCTCTTAAACAGCTAGCAACGTAAC 992
Db      327 AspValAsnThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
Qy      993 TTAGACTTCAGACATTTATACGATCTCTCGTGATAAGCTAAACTTACTCTACAACAATCTC 1052
Db      347 LeuAspPheArgAspLeuTyArgProArgAspLysAlaLysLeuLeuTyAsnAsnLeu 366
Qy      1053 GATGCTTTTGGTATTATGGAATACCTTAACTGGAAGCTAGAGGATATATCAGTATGAC 1112
Db      367 AspAlaPheGlyIleMetAspTyThrLeuThrGlyLysValGluAspAsnHisAspAsp 386
Qy      1113 ACCAACCGTATATAACCGTTTATATGGGCAAGCGACCCGAAGGAGAGATGCTAGCTAC 1172
Db      387 ThrAsnArgIleIleThrValTyMetGlyLysArgProGluGlyGluAsnAlaSerTy 406
Qy      1173 CATTAGCT 1181
Db      407 HisLeuAla 409
```

Db	247	SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu	266
Qy	753	TTTACTTACCGTGTAAATAATCGGAAACAAGCTTATAGAGTCAATAAAAAATCTGGTCTG	812
Db	267	PheThrTyrHisValIysAsnArgGluGlnAlaTyrGluIleAsnProLysThrGlyIle	286
Qy	813	AATGAAGAAATAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG	872
Db	287	LysGluLysThrAsnAsnThrAspLeuValSerGluLysTyrTyrValLeuLysGlnGly	306
Qy	873	GAAGAAGCGTATGATCCCTTTGATCGCAGTCACTTGAACCTTCACCATCAATACGTT	932
Db	307	GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal	326
Qy	933	GATGTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAACAGCTAGCGAAGCTAAC	992
Db	327	AspValAsnThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluGlyAsn	346
Qy	993	TTAGACTTCAGAGATTTATACGATCTCTCGTGATAAGGCTAAACTACTCTACAACAATCTC	1052
Db	347	LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu	366
Qy	1053	GATGCTTTTCGGTATTATGGAATATACCTTAACCTGGAAGAGTAGAGGATATATCAGATGAC	1112
Db	367	AspAlaPheAspIleMetAsnTyrThrLeuThrGlyLysValGluAspAsnHisAspLys	386
Qy	1113	ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAGAGAGAGAATGCTAGCTAC	1172
Db	387	AsnAsnArgValValThrValTyrMetGlyLysArgProLysGlyAlaLysGlySerTyr	406
Qy	1173	CATTTAGCT 1181	
Db	407	HisLeuAla 409	
RESULT 5			
JU0292			
streptokinase - Streptococcus pyogenes			
C:Species: Streptococcus pyogenes			
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004			
C:Accession: A43867; JU0292			
R:Okuni, H.; Todome, Y.; Suzuki, H.; Mizuse, M.; Kotani, N.; Horiuchi, K.; Shi			
A:Title: Immunochemical studies and complete amino acid sequence of the streptol			
A:Reference number: A43867; MUID:92104686; PMID:1370275			
A:Accession: A43867			
A:Molecule type: protein			
A:Residues: 1-414 <OHK>			
A:Cross-references: UNIPROT:Q57391; UNIPARC:UPI0000175C82			
A:Experimental source: M type 12 strain A374			
A:Note: sequence extracted from NCBI backbone (NCBIP:74592)			
C:Superfamily: streptokinase			
Alignment Scores:			
Pred. No.:	3,56e-106	Length:	414
Score:	1673.00	Matches:	324
Percent Similarity:	90.60%	Conservative:	23
Best Local Similarity:	84.60%	Mismatches:	36
Query Match:	61.60%	Indels:	0
DB:	2	Gaps:	0
US-09-940-235-9 (1-1541) x JU0292 (1-414)			
Qy	33	ATAGCTGGTCTCGAATGGCTACTAGATCGTCTCTGTAATAATACACCAATGGTTGTT	92
Db	1	IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal	20
Qy	93	AGCGTGTCTGGTACTGTTGAGGGGCAATCAAGACATAGTCTTAAATTTTTTGAATC	152
Db	21	SerValAlaGlyThrValGluGlyThrAsnGlnGluIleSerLeuLysPhePheGluIle	40
Qy	153	GATCTAACATCAGCACTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAGTCCAAATCA	212

Db 41 AspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 60
QY 213 AAACATTTGCTACTGATGTCGCGGATGTCATATAAATTCGAGAAGCTGACTACTA 272
Db 61 LysProPheAlaThrAsnSerSerAlaMetProHisLysLeuGluLysAlaAspLeu 80
QY 273 AAGGCTATTCAAGAACAAATTCGTTAAGTCCACAGTACACGACGACTACTTTGAGGTC 332
Db 81 LysAlaIleGlnGluLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal 100
QY 333 ATTGATTTTGCAGCGATCAACCATTTACTGATCGAAACCGGAAAGTCTACTTTGCTGAC 392
Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
QY 393 AAGATGTTTCGTTACCTTCGCGACCAACCTGTCACAAATTTTTCGTTAAGCGACAT 452
Db 121 ArgAspAspSerValThrLeuProThrGlnProValGlnGluPheLeuSerGlyHis 140
QY 453 GTGCGGTTAGCCATATATAAGAAACCAATACAAACCAACGAAATCTGTTGATGTG 512
Db 141 ValArgValArgProTyrGlnProLysAlaValHisAsnSerAlaGluArgValAsnVal 160
QY 513 GAATATAGTGTACAGTTTACTCCCTTAACCCCTGATGACGATTTTCAGACGAGTCTCAA 572
Db 161 AsnTyrGluValSerPheValSerGluThrGlyAsnLeuAspPheThrProSerLeuLys 180
QY 573 GATACCTAAGCTATTGAACACATAGCTATCGGTGACACCATCATCATCAAGAACTACTA 632
Db 181 GluArgTyrHisLeuThrThrLeuAlaValAlaGlyAspSerLeuSerGlnGluLeuAla 200
QY 633 GCTCAAGCACAAAGCATTTTAAACCAAAACCCAGGCTATACGATTTATGAAACGTGAC 692
Db 201 AlaIleAlaGlnPheIleLeuSerLysGluHisProAspTyrIleIleThrLysArgAsp 220
QY 693 TCCTCAATGCTACTGACAAATGACATTTCCGTCAGATTTTACCAATGGATCAAGAG 752
Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
QY 753 TTTACTTACCGTGTAAATATCGGACGAGTATAGCTTATAGATCATATAAATCTGCTCG 812
Db 241 PheThrTyrHisIleLysAspArgGluGlnAlaTyrLysAlaAsnSerLysThrGlyIle 260
QY 813 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGCTCTTAAAGAGG 872
Db 261 ValGluLysThrAsnAsnThrAspLeuIleSerGluLysTyrValLeuLysLysGly 280
QY 873 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTTCAACCATCAATACGTT 932
Db 281 GluGluProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
QY 933 GATGTCGATACCAACGAATGCTAAGAGTGAAGCTCTTAAACAGCTAGCGAACGTAAC 992
Db 301 AspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320
QY 993 TTGACTTCAGAGATTATACCATCTCTGATAGGCTTAAGCTTAACTCTTCAACATCTC 1052
Db 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
QY 1053 GATGCTTTGTTATGACTATACCTTAACTGGAAGATGAGGATATACAGATGAC 1112
Db 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAsnAsp 360
QY 1113 ACCAACCGTATCAATACCGTTTATATGGCAAGCGACCCGAGGAGGAGATGCTAGCTAC 1172
Db 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 380
QY 1173 CATTTAGCT 1181
Db 381 HisLeuAla 383
RESULT 6
S04168
streptokinase A precursor - Streptococcus pyogenes (strain NZ131)

C;Species: Streptococcus pyogenes
C;Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 09-Jul-2004
C;Accession: S04168
R;Huang, T.T.; Malke, H.; Ferretti, J.J.
Mol. Microbiol. 3, 197-205, 1989
A;Title: The streptokinase gene of group A streptococci: cloning, expression in Escheri-
A;Reference number: S04168; MUID:89343623; PMID:2668686
A;Accession: S04168
A;Molecule type: DNA
A;Residues: 1-440 <HUA>
A;Cross-references: UNIPROT:Q57391; UNIPARC:UPI00000BD04A; EMBL:X51517; NID:g47437; PID:
C;Genetics:
A;Gene: ska
C;Superfamily: streptokinase
F;1-26/Domain: signal sequence #status predicted <SIG>
F;27-440/Product: streptokinase A #status predicted <MAT>
Alignment Scores:
Pred. No.: 5,73e-106 Length: 440
Score: 1670.00 Matches: 324
Percent Similarity: 90.34% Conservative: 22
Best Local Similarity: 84.60% Mismatches: 37
Query Match: 61.49% Indels: 0
DB: 2 Gaps: 0
US-09-940-235-9 (1-1541) x S04168 (1-440)
QY 33 ATAGCTGGTCTCTGAATGCTACTAGATCGTCCTTCTCTTAATAACAGCCAAATCGTGT 92
Db 27 IleAlaGlyTyrGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 46
QY 93 AGCGTCTGCTACTGTTGAGGCGCAATCAAGACATTAAGTCTTAATTTTGAATC 152
Db 47 SerValAlaGlyThrValGluGlyThrAsnGlnGluIleSerLeuLysPhePheGluIle 66
QY 153 GATCTAATCATCACACCTGCTCATGGAGGAGAAAGACAGACGAGCTTAAGTCCAAATCA 212
Db 67 AspLeuThrSerArgProAlaHisGlyGlyThrGlnGlnGlyLeuSerProLysSer 86
QY 213 AAACATTTGCTACTGATGTCGCGGATGTCATATAAATTCGAGAAGCTGACTACTA 272
Db 87 LysProPheAlaThrAspLysGlyAlaMetSerHisLysLeuGluLysAlaAspLeu 106
QY 273 AAGGCTATTCAAGAACAAATTCGTCGCTTAACGTCACAGTAAACGACGACTACTTTGAGGTC 332
Db 107 LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal 126
QY 333 ATTGATTTTGCAGCGATGCAACCATTTACTGATCGAAACCGGAAAGTCTACTTTGCTGAC 392
Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146
QY 393 AAGATGTTTCGTTACCTTCGCGACCAACCTGTCACAAATTTTTCGTTAAGCGACAT 452
Db 147 LysAspAspSerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 166
QY 453 GTGCGGCTTAGACATATAAGAAACCAATACAAACCAACGAGGAAATCTGTTGATGTG 512
Db 167 ValArgValLysProTyrGlnProLysAlaValHisAsnSerAlaGluArgValAsnVal 186
QY 513 GAATATAGTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTCAA 572
Db 187 AsnTyrGluValSerPheValSerGluThrGlyAspLeuAspPheThrProLeuLeuArg 206
QY 573 GATACCTAAGCTATTGAACACATAGCTATCGGTGACACCATCATCATCAAGAAATACTA 632
Db 207 AsnGlnTyrHisLeuThrThrLeuAlaValGlyAspSerLeuSerGlnGluLeuAla 226
QY 633 GCTCAAGCACAAAGCATTTTAAACCAAAACCCAGGCTATACGATTTATGAAACGTGAC 692
Db 227 AlaIleAlaGlnPheIleLeuSerLysLysHisProAspTyrIleIleThrLysArgAsp 246
QY 693 TCCTCAATGCTACTGATGACAAATGACATTTTCGTCAGTATGATTTTACCAATGGATCAAGAG 752

Db 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
 QY 753 TTTACTACCGTGTAAATAATCGGAACAAGCTTATAGGATCAATATAAAAAATCTGCTCG 812
 Db 267 PheThrThrHisIleLysAspArgGluGlnAlaTyrLysAlaAsnSerLysThrGlyIle 286
 QY 813 AATGAAGAAATAACAACACACCTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 872
 Db 287 GluGluLysThrAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 306
 QY 873 GAAAAGCGTATGATCCCTTTGATCGAGTCACTTGAAGAACTGTTTCAACATCAATACGTT 932
 Db 307 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleAsnTyrVal 326
 QY 933 GATGTCGATACCAAGAAATTGCTAAAGTAGAGCTCTTAACAGCTAGCGAAGCTAAC 992
 Db 327 AspValAsnThrAsnLysLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
 QY 993 TTAGACTTCAGAGATTATACGATCTCTGTGATAAGGCTTAAACTACTTACACCAATCTC 1052
 Db 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
 QY 1053 GATGCTTTGGTATTATGACATATACCTTAACTGGAAGTAGAGGATTAATCAAGATGAC 1112
 Db 367 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 386
 QY 1113 ACCAAGCTATCATACCGTTTATATGGCAGACGACCGAAGGAGAGATGCTAGCTAC 1172
 Db 387 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 406
 QY 1173 CATTTAGCT 1181
 Db 407 HisLeuAla 409

RESULT 7

S53334
 streptokinase - Streptococcus sp.
 C;Species: Streptococcus sp.
 C;Date: 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change 09-Jul-2004
 R;Shi, G.Y.; Chang, B.I.; Chen, S.M.; Wu, D.H.; Wu, H.L.
 Biochem. J. 304, 235-241, 1994
 A;Title: Function of streptokinase fragments in plasminogen activation.
 A;Reference number: S53334; MUID:95091634; PMID:7998939
 A;Accession: S53334
 A;Status: preliminary
 A;Molecule type: protein
 A;Residues: 1-197 <SH1>
 A;Cross-references: UNIPROT:Q7M115; UNIPARC:UPI000017AC2C

Alignment Scores:

Pred. No.:	8.69e-38	Length:	197
Score:	667.00	Matches:	167
Percent Similarity:	44.71%	Conservative:	2
Best Local Similarity:	44.18%	Mismatches:	1
Query Match:	24.56%	Indels:	208
DB:	2	Gaps:	9

US-09-940-235-9 (1-1541) x S53334 (1-197)

QY 33 ATAGCTGTCTGAATGGCTACTAGATCGTCTCTGTAAATAACAGCCAATTGGTTGTT 92
 Db 1 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsn----- 14
 QY 93 AGCGTTGTGCTACTGTGTTGAGGGGAGCAATCAAGCAATTAGTCTTAAATTTTGGAAATC 152
 Db 14 ----- 14
 QY 153 GATCTTAACATCAAGACCTGCTCATGGAGAAAGACAGACGAGCTTAAGTCAAAATCA 212
 Db 15 -----Ser 15
 QY 213 AAACCAATTTGCTACTGATAGTGGCGGATGTACATAACTTGAGAAAGCTGACTACTA 272

Db 16 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 35
 QY 273 AAGGCTATTCAAGAACAAATTGATCGTAACTCCACAGTACACGACGACTACTTTGAGTTC 332
 Db 36 LysAlaIleGlnTrpGlnIleLeu----- 43
 QY 333 ATTGATTTTGAAGCGATGCAACCATTTACTGATCGAAACGGCAAGGCTCTACTTTGCTGAC 392
 Db 44 -----AsnGlyLysValTyrPheAlaAsp 51
 QY 393 AAAGATGGTTCGGTAACTTTGCGACCAACCTGTCCAAGANTTTTGTCTAAGCGGCAT 452
 Db 52 LysAspGlySerValThr----- 57
 QY 453 GTGCGGTTAGACCATATAAAGAAAAACCAATACAAAACCAAGGAAATCTGTTGATGTG 512
 Db 58 -----GluLysProIleGlnAsnGlnAlaLysSerValAspVal 70
 QY 513 GAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCAGGTCTCAA 572
 Db 71 Glu----- 71
 QY 573 GATACTAAGCTATTGAAAACACACTAGCTATCGGTGACACCATCATCTCAAGAATTACTA 632
 Db 71 ----- 71
 QY 633 GCTCAAGCAACAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTATGAACGTCG 692
 Db 72 -----AsnHisProGlyTyrThrIleTyrGluArgAsp 82
 QY 693 TCCTCAATCGTCACTCATGACATGACATTTTCGGTACGATTTTACCAATGGATCAAGAG 752
 Db 83 SerSerIleVal-----ThrIleLeuProMetAspGlnGlu 94
 QY 753 TTTTACTTACCGTGTGTAAATAATCGGAACAAGCTTATAGGATCAATAAAAAATCTGCTCTG 812
 Db 95 PheThrTyr-----SerGlyLeu 100
 QY 813 AATGAAGAAATAACAACACTGACTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 872
 Db 101 AsnGluGluIleAsnAsnThrAspLeuIleSer----- 111
 QY 873 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCCACCATCAATACGTT 932
 Db 112 -----TyrVal 113
 QY 933 GATGTCGATACCAACGAATTGTCTAAAGTAGACAGCTCTTAAACAGCTAGCAACGTAAC 992
 Db 114 AspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 133
 QY 993 TTAGACTTCAGAGATTATACGATCTCTGTCGTAAGCTTAACTACTCTACACAATCTC 1052
 Db 134 -----AspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 149
 QY 1053 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAGTAGAGGATTAATCAGATGAC 1112
 Db 150 AspAlaPheGlyIleMet----- 155
 QY 1113 ACCAAGCTATCATACCGTTTATATGGCAAGCGACCCGAGGAGAGAAATGCT 1166
 Db 156 -----IleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAla 170

RESULT 8

FNHU

fibronectin precursor [validated] - human

N;Alternate names: fibronectin splice form ED-A

C;Species: Homo sapiens (man)

C;Date: 27-Nov-1985 #sequence_revision 31-Mar-1993 #text_change 31-Dec-2004

C;Accession: A26460; A26284; S03917; A24854; A24476; A91008; A93529; A21011; A90495; A32

R;Dean, D.C.; Bowlus, C.L.; Bourgeois, S.

Proc. Natl. Acad. Sci. U.S.A. 84, 1876-1880, 1987

A;Title: Cloning and analysis of the promoter region of the human fibronectin gene.

A;Reference number: A26460; MUID:87175578; PMID:3031656
 A;Accession: A26460
 A;Molecule type: DNA
 A;Residues: 1-49 <DEA>
 A;Cross-references: UNIPROT:P02751; UNIPROT:Q14327; UNIPARC:UPI000016A926; GB:M15801; NID:1494
 R;Oldberg, A.; Ruoslahti, E.
 J. Biol. Chem. 261, 2113-2116, 1986
 A;Title: Evolution of the fibronectin gene.
 A;Reference number: A26284; MUID:86111901; PMID:3003095
 A;Accession: A26284
 A;Molecule type: DNA
 A;Residues: 1447-1540 <OLD>
 A;Cross-references: UNIPARC:UPI0000112E37; GB:M12549; NID:g182688
 A;Note: the authors translated the codon TTC for residue 1494 as Glu
 R;Paolella, G.; Henchcliffe, C.; Sebastio, G.; Baralle, F.E.
 Nucleic Acids Res. 16, 3545-3557, 1988
 A;Title: Sequence analysis and in vivo expression show that alternative splicing of ED-E
 A;Reference number: S00848; MUID:88233940; PMID:3375063
 A;Accession: S03917
 A;Molecule type: DNA
 A;Residues: 1594-1767, 'V', 1769-1783 <PAO>
 A;Cross-references: UNIPARC:UPI000017432B; EMBL:X07718; NID:g31402
 A;Note: the authors translated the codon AAC for residue 1631 as Asp
 R;Vibe-Pedersen, K.; Magnusson, S.; Baralle, F.E.
 FEBS Lett. 207, 287-291, 1986
 A;Title: Donor and acceptor splice signals within an exon of the human fibronectin gene.
 A;Reference number: A24854; MUID:87030929; PMID:3770201
 A;Accession: A24854
 A;Molecule type: DNA
 A;Residues: 1992-2147 <VIB>
 A;Cross-references: UNIPARC:UPI000017432C; GB:X04530; NID:g31436
 R;Gutman, A.; Yamada, K.M.; Kornblihtt, A.
 FEBS Lett. 207, 145-148, 1986
 A;Title: Human fibronectin is synthesized as a pre-propolypeptide.
 A;Reference number: A24476; MUID:87030890; PMID:3770189
 A;Accession: A24476
 A;Status: not compared with conceptual translation
 A;Molecule type: mRNA
 A;Residues: 1-14, 'Q', 16-38 <GUT>
 A;Cross-references: UNIPARC:UPI000017432D
 R;Kornblihtt, A.R.; Umezawa, K.; Vibe-Pedersen, K.; Baralle, F.E.
 EMBO J. 4, 1755-1759, 1985
 A;Title: Primary structure of human fibronectin: differential splicing may generate at l
 A;Reference number: A91008; MUID:85284965; PMID:2992939
 A;Accession: A91008
 A;Status: nucleic acid sequence not shown
 A;Molecule type: mRNA
 A;Residues: 32-1344, 1346-2080; 2112-2386 <KOR>
 A;Cross-references: UNIPARC:UPI000017432E; UNIPARC:UPI000017432F; GB:X02761
 R;Kornblihtt, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
 Nucleic Acids Res. 12, 5853-5868, 1984
 A;Title: Human fibronectin: cell specific alternative mRNA splicing generates polypeptid
 A;Reference number: A93529; MUID:84272258; PMID:6462919
 A;Accession: A93529
 A;Molecule type: mRNA
 A;Residues: 973-2080; 2112-2386 <K2>
 A;Cross-references: UNIPARC:UPI000017432F; UNIPARC:UPI0000174330; GB:X00739
 R;Oldberg, A.; Linney, E.; Ruoslahti, E.
 J. Biol. Chem. 258, 10193-10196, 1983
 A;Title: Molecular cloning and nucleotide sequence of a cDNA clone coding for the cell a
 A;Reference number: A21011; MUID:83290929; PMID:6688418
 A;Accession: A21011
 A;Molecule type: mRNA
 A;Residues: 1434-1537 <OL2>
 A;Cross-references: UNIPARC:UPI0000174331; GB:X00055; NID:g182680; PIDN:AAA52459.1; PID:1
 R;Bernard, M.P.; Kolbe, M.; Weil, D.; Chu, M.L.
 Biochemistry 24, 2698-2704, 1985
 A;Title: Human cellular fibronectin: comparison of the carboxyl-terminal portion with ra
 A;Reference number: A90495; MUID:85280409; PMID:2992573
 A;Accession: A90495
 A;Molecule type: mRNA
 A;Residues: 1594-2386 <BER>
 A;Cross-references: UNIPARC:UPI000016A928; GB:M10905; NID:g182696; PIDN:AAA52462.1; PID:1

R;Umezawa, K.; Kornblihtt, A.R.; Baralle, F.E.
 FEBS Lett. 186, 31-34, 1985
 A;Title: Isolation and characterization of cDNA clones for human liver fibronectin.
 A;Reference number: A22245; MUID:85231203; PMID:2989004
 A;Accession: A22245
 A;Molecule type: mRNA
 A;Residues: 1948-2067 <UME>
 A;Cross-references: UNIPARC:UPI000000046A; GB:M27589; NID:g182705; PIDN:AAA52465.1; PID:1
 A;Accession: B22245
 A;Molecule type: mRNA
 A;Residues: 1975-1991, 2017-2039 <UM2>
 A;Cross-references: UNIPARC:UPI0000174332; UNIPARC:UPI0000174333; GB:M27590
 R;Sekiguchi, K.; Klos, A.M.; Kurachi, K.; Yoshitake, S.; Hakomori, S.
 Biochemistry 25, 4936-4941, 1986
 A;Title: Human liver fibronectin complementary DNAs: identification of two different me
 A;Reference number: 152394; MUID:87026578; PMID:3021206
 A;Accession: 165273
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1978-1990, 2016-2018, 'N', 2020-2081, 2113-2127 <SEK>
 A;Cross-references: UNIPARC:UPI000006E04C; GB:M14060; NID:g182701; PIDN:AAA52464.1; PID:1
 R;Kornblihtt, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
 Proc. Natl. Acad. Sci. U.S.A. 80, 3218-3222, 1983
 A;Title: Isolation and characterization of cDNA clones for human and bovine fibronectin
 A;Reference number: A21165; MUID:83221567; PMID:6304699
 A;Accession: A21165
 A;Molecule type: mRNA
 A;Residues: 2291-2386 <K03>
 A;Cross-references: UNIPARC:UPI0000174334; GB:X00799; NID:g182681; PIDN:AAA52460.1; PID:1
 R;Garcia-Pardo, A.; Pearlstein, E.; Frangione, B.
 J. Biol. Chem. 258, 12670-12674, 1983
 A;Title: Primary structure of human plasma fibronectin.
 A;Reference number: A92398; MUID:84032463; PMID:6630202
 A;Accession: A92398
 A;Molecule type: protein
 A;Residues: 32-47, 'C', 49-51, 'S', 53-72, 'A', 74-290 <GAR1>
 A;Cross-references: UNIPARC:UPI0000174335
 R;Garcia-Pardo, A.; Gold, L.I.
 Arch. Biochem. Biophys. 304, 181-188, 1993
 A;Title: Further characterization of the binding of fibronectin to gelatin reveals the p
 A;Reference number: S34791; MUID:93312001; PMID:8323285
 A;Accession: S34791
 A;Molecule type: protein
 A;Residues: 291-300, 551-560 <GAR2>
 A;Cross-references: UNIPARC:UPI0000174336; UNIPARC:UPI0000174337
 R;Griffin, C.A.; Calaycay, J.; Shively, J.E.; Smith, R.L.
 Thromb. Res. 43, 469-477, 1986
 A;Title: Two plasma fibronectin fragments with different gelatin-binding properties.
 A;Reference number: A60904; MUID:87019725; PMID:3532418
 A;Accession: A60904
 A;Molecule type: protein
 A;Residues: 293-301 <GRI>
 A;Cross-references: UNIPARC:UPI0000174338
 R;Calaycay, J.; Pande, H.; Lee, T.; Borsai, L.; Siri, A.; Shively, J.E.; Zardi, L.
 J. Biol. Chem. 260, 12136-12141, 1985
 A;Title: Primary structure of a DNA- and heparin-binding domain (domain III) in human p
 A;Reference number: A23901; MUID:86008277; PMID:3900070
 A;Accession: A23901
 A;Molecule type: protein
 A;Residues: 616-677, 'Q', 679-703, 'PT' <CAL>
 A;Cross-references: UNIPARC:UPI0000174339
 R;Pierschbacher, M.D.; Ruoslahti, E.; Sundelin, J.; Lind, P.; Peterson, P.A.
 J. Biol. Chem. 257, 9593-9597, 1982
 A;Title: The cell attachment domain of fibronectin. Determination of the primary structu
 A;Reference number: A92386; MUID:82265604; PMID:7050098
 A;Accession: A92386
 A;Molecule type: protein
 A;Residues: 1441-1548 <PIE>
 A;Cross-references: UNIPARC:UPI0000141CD5
 A;Note: residues 1524-1527 are responsible for the cell-binding activity
 R;Garcia-Pardo, A.; Rostagno, A.; Frangione, B.
 Biochem. J. 241, 923-928, 1987
 A;Title: Primary structure of human plasma fibronectin. Characterization of a 38 kDa dom


```

Db      101 PheThrTyHisValysAsnArgGluGlnAlaTyRGluIleAsnProIysThrGlyIle 120
QY      813 AATGAAGAAATAAACCAACTGAC 836
Db      121 LysGluLysThrAsnAsnThrAsp 128

RESULT 10
FNBO
fibronectin - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 31-Dec-1988 #sequence revision 31-Dec-1988 #text_change 09-Jul-2004
C:Accession: A26452; B21165; A23292
R:Skorstengard, K.; Jensen, M.S.; Sahli, P.; Petersen, T.E.; Magnusson, S.
Eur. J. Biochem. 161, 441-453, 1986
A:Title: Complete primary structure of bovine plasma fibronectin.
A:Reference number: A26452; MUID:87054047; PMID:3780752
A:Accession: A26452
A:Molecule type: protein
A:Residues: 1-2265 <SKO>
A:Cross-references: UNIPROT:P07589; UNIPARC:UPI000012A7BE
R:Kornblihtt, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
Proc. Natl. Acad. Sci. U.S.A. 80, 3218-3222, 1983
A:Title: Isolation and characterization of cDNA clones for human and bovine fibronectin
A:Reference number: A21165; MUID:83221567; PMID:6304699
A:Accession: B21165
A:Molecule type: mRNA
A:Residues: 2170-2265 <KOR>
A:Cross-references: UNIPARC:UPI000016C30C; GB:K00800; NID:G163055; PIDN:AAA30521.2; PID:
R:Petersen, T.E.; Thøgersen, H.C.; Skorstengard, K.; Vibe-Pedersen, K.; Sahli, P.; Sottr
Proc. Natl. Acad. Sci. U.S.A. 80, 137-141, 1983
A:Title: Partial primary structure of bovine plasma fibronectin: three types of internal
A:Reference number: A23292; MUID:83117805; PMID:6218503
A:Accession: A23292
A:Molecule type: protein
A:Residues: 1-16, 'C', 18-20, 'S', 22-432, 447-463, 1367-1517, 1567-1673; 2062-2176, 'N', 2178-226
A:Cross-references: UNIPARC:UPI000017433E; UNIPARC:UPI000017433F; UNIPARC:UPI0000174340;
C:Comment: Cys-1201 and Cys-2015 have free sulphydryl groups.
C:Comment: The plasma fibronectin molecule consists of two chains, which are connected b
C:Comment: Fibronectins bind cell surfaces and various compounds including collagen, fib
aling, and maintenance of cell shape.
C:Comment: Plasma fibronectin is synthesized by hepatocytes.
C:Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II rep
C:Keywords: acute phase; alternative splicing; collagen binding; duplication; extracellu
F:21-241/Domain: fibrin and heparin binding <FBR>
F:21-56/Domain: fibronectin type I repeat homology <1F1>
F:66-104/Domain: fibronectin type I repeat homology <1F2>
F:110-148/Domain: fibronectin type I repeat homology <1F3>
F:155-194/Domain: fibronectin type I repeat homology <1F4>
F:200-239/Domain: fibronectin type I repeat homology <1F5>
F:277-577/Domain: collagen binding <CBR>
F:277-311/Domain: fibronectin type I repeat homology <1F6>
F:329-370/Domain: fibronectin type II repeat homology <2F1>
F:389-430/Domain: fibronectin type II repeat homology <2F2>
F:439-477/Domain: fibronectin type I repeat homology <1F7>
F:487-524/Domain: fibronectin type I repeat homology <1F8>
F:530-568/Domain: fibronectin type I repeat homology <1F9>
F:578-661/Domain: fibronectin type III repeat homology <FN3A>
F:688-770/Domain: fibronectin type III repeat homology <FN3B>
F:779-860/Domain: fibronectin type III repeat homology <FN3C>
F:875-957/Domain: fibronectin type III repeat homology <FN3D>
F:965-1046/Domain: fibronectin type III repeat homology <FN3E>
F:1055-1134/Domain: fibronectin type III repeat homology <FN3F>
F:1142-1227/Domain: fibronectin type III repeat homology <FN3G>
F:1235-1318/Domain: fibronectin type III repeat homology <FN3H>
F:1326-1404/Domain: fibronectin type III repeat homology <GN3I>
F:1410-1517/Domain: cell attachment <CAD>
F:1416-1502/Domain: fibronectin type III repeat homology <FN3J>
F:1493-1495/Region: cell attachment (R-G-D) motif
F:1510-1592/Domain: fibronectin type III repeat homology <FN3K>
F:1600-1870/Domain: heparin binding <HB2>
F:1600-1682/Domain: fibronectin type III repeat homology <FN3L>
F:1692-1773/Domain: fibronectin type III repeat homology <FN3M>

```

```

F:1781-1863/Domain: fibronectin type III repeat homology <FN3N>
F:1970-1972/Region: cell attachment (R-G-D) motif
F:1982-2062/Domain: fibronectin type III repeat homology <FN3O>
F:1985-2216/Domain: fibrin binding <FB2>
F:2085-2124/Domain: fibronectin type I repeat homology <1F10>
F:2130-2167/Domain: fibronectin type I repeat homology <1F11>
F:2174-2209/Domain: fibronectin type I repeat homology <1F12>
F:1/Modified site: pyroliidone carboxylic acid (Gln) #status experimental
F:3/Cross-link: isopeptide (Gln) (interchain to Lys N6-amino of fibrin) #status experimental
F:21-47, 45-56, 66-94, 92-104, 110-138, 138-148, 155-184, 182-194, 200-229, 227-239, 277-304, 302-
7, 2155-2167, 2174-2200, 2198-2209/Disulfide bonds: #status predicted
F:399, 497, 511, 846, 976, 1213, 1987/Binding site: carbohydrate (Asn) (covalent) #status absent
F:1205, 1692/Binding site: carbohydrate (Asn) (covalent) #status experimental
F:1243, 1944/Binding site: carbohydrate (Thr) (covalent) #status experimental
F:2246/Disulfide bonds: interchain (to 2250) #status predicted
F:2250/Disulfide bonds: interchain (to 2246) #status predicted
F:2263/Binding site: phosphate (Ser) (covalent) #status experimental

Alignment Scores:
Pred. No.:      8.14e-34      Length:      2265
Score:          610.00      Matches:    113
Percent Similarity: 86.13%      Conservative: 5
Best Local Similarity: 82.48%      Mismatches: 13
Query Match:    22.46%      Indels:      6
DB:             1          Gaps:        3

US-09-940-235-9 (1-1541) x FNBO (1-2265)
QY      1143 AAGCGACCCGAGGAGAGATGCTAGTACCATTTA-----GCTGGTGGTGGC 1190
Db      124 ArgArgProHis---GluThrGlyGlyTyMetLeuGluCysValCysLeuGlyAsnGly 142
QY      1191 CAGCGCGAA---CAGATTGTACCATAGCTGAGAGAGTGTTTGATCATGCTGCTGGGACT 1247
Db      143 LysGlyGlnTrpThrCysLysProIleAlaGluLysCysPheAspGlnAlaAlaGlyThr 162
QY      1248 TCCTATGTGTCGGAGAAACCTGGGCAAGCCCTACCAAGCTCGATGATGTTAGATTGT 1307
Db      163 SerTyrValValGlyGluThrTrpGluLysProIleGlnGlyTrpMetMetValAspCys 182
QY      1308 ACTTGTCTGGGAGAGGACGACGACGATCATCTTGACATCTTCTAGAAATAGATGCAACGAT 1367
Db      183 ThrCysLeuGlyGluGlySerGlyArgIleThrCysThrSerArgAsnArgCysAsnAsp 202
QY      1368 CAGGACACAGGACATCTTAGAATTGGAGACACCTGGAGCAAGAGAGGATAATCGAGGA 1427
Db      203 GlnAspThrArgThrSerTyrArgIleGlyAspThrTrpSerLysLysAspAsnArgGly 222
QY      1428 AACCTGCTCCAGTGCATCTGCACAGGCAACGGCGGAGAGAGTGGAGTCTGAGAGGCAC 1487
Db      223 AsnLeuLeuGlnCysIleCysThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHis 242
QY      1488 ACCTCTGTGACGACCATCGAGCGGATCTGGCCCTTCCACCGATGTTCTG 1538
Db      243 ThrSerLeuGlnThrThrSerAlaGlySerGlySerPheThrAspValArg 259

RESULT 11
S14428
fibronectin precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 28-Oct-1994 #sequence revision 28-Oct-1994 #text_change 09-Jul-2004
C:Accession: S14428; S12455; A22319; S46203; S00459; A27252; I59049
R:Hynes, R.O.
submitted to the EMBL Data Library, July 1989
A:Reference number: S14428
A:Accession: S14428
A:Molecule type: mRNA
A:Residues: 1-2477 <HYN>
A:Cross-references: UNIPROT:P04937; UNIPARC:UPI000012A7C6; EMBL:X15906; NID:G56163; PIDN:
R:Schwarzbauer, J.E.; Patel, R.S.; Fonda, D.; Hynes, R.O.
EMBO J. 6, 2573-2580, 1987
A:Title: Multiple sites of alternative splicing of the rat fibronectin gene transcript.
A:Reference number: S12455; MUID:88054951; PMID:2445560

```


F;1356-1439/Domain: fibronectin type III repeat homology <FN3I>
F;1447-1529/Domain: fibronectin type III repeat homology <FN3J>
F;1537-1619/Domain: fibronectin type III repeat homology <FN3K>
F;1614-1616/Region: cell attachment (R-G-D) motif
F;1631-1713/Domain: fibronectin type III repeat homology <FN3L>
F;1721-1803/Domain: fibronectin type III repeat homology <FN3M>
F;1811-1893/Domain: fibronectin type III repeat homology <FN3N>
F;1903-1984/Domain: fibronectin type III repeat homology <FN3O>
F;1992-2074/Domain: fibronectin type III repeat homology <FN3P>
F;2181-2183/Region: cell attachment (R-G-D) motif
F;2193-2273/Domain: fibronectin type III repeat homology <FN3Q>
F;2296-2335/Domain: fibronectin type I repeat homology <IF10>
F;2341-2378/Domain: fibronectin type I repeat homology <IF11>
F;2385-2420/Domain: fibronectin type I repeat homology <IF12>
F;53-79,77-88,126-136,142-170,168-180,187-216,214-226,232-261,259-271,308-335,333-368,2366-2378,2385-2411,2409-2420/Disulfide bonds: #status predicted
F;2458/Disulfide bonds: interchain (to 2462) #status predicted
F;2462/Disulfide bonds: interchain (to 2458) #status predicted

Alignment Scores:
Pred. No.: 3.8e-32 Length: 2477
Score: 585.50 Matches: 108
Percent Similarity: 84.67% Conservative: 8
Best Local Similarity: 78.83% Mismatches: 14
Query Match: 21.56% Indels: 7
DB: 2 Gaps: 4

US-09-940-235-9 (1-1541) x S14428 (1-2477)

Qy 1143 AAGCGACCGAAGAGAGAGATGCTAGCTACCATTTA-----GCTGGTGGTGGC 1190
Db 156 ArgArgProHis---GluThrGlyGlyTyrMetLeuGluCysLeuCysLeuGlyAsnGly 174
Qy 1191 CAGCGCAA---CAGATTGTACCCATAGCTGAGAGAGTGTTCATCATGCTCGTGGCACT 1247
Db 175 LysGlyGluTrpThrCysLysProIleAlaGluLysCysPheAspHisAlaIaGlyThr 194
Qy 1248 TCCTATGTGTCGGAGAAACCTGGAGAAGCCCTACCAAGCTCGATGCTAGATTGT 1307
Db 195 SerTyrValValGlyGluThrTrpGluLysProTyrGlnGlyTrpMetMetValAspCys 214
Qy 1308 ACTTGCTGGAGAAGCAGCGACGCATCATTGCACTTCTAGAAATAGATGCAACGAT 1367
Db 215 ThrCysLeuGlyGluGlyAsnGlyArgIleThrCysThrSerArgAsnArgCysAsnAsp 234
Qy 1368 CAGACACAGGACATCCTATAGATTGGAGACACCTGGACGACGAGAGAGGATATCGAGGA 1427
Db 235 GlnAspThrArgThrSerTyrArgIleGlyAspThrTrpSerLysLysAspAsnArgGly 254
Qy 1428 AACCTGCTCCAGTCATCTGCACAGGCACACGGCCGAGAGAGTGGAAAGTGTGAGAGCAC 1487
Db 255 AsnLeuLeuGlnCysValCysThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHis 274
Qy 1488 ACTCTCTGCAGACCATCAGCGGATCTGGCCCTTCACCGGATTCGT 1538
Db 275 Val---LeuGlnSerAlaSerAlaGlySerGlySerPheThrAspValArg 290

RESULT 12
S77680
streptokinase A (SC 3.4.-.-) (allele 11) - Streptococcus pyogenes (fragment)
C:Species: Streptococcus pyogenes
C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004
C:Accession: S77680
R:Kapur, V.; Kanjilal, S.; Hamrick, M.R.; Li, L.L.; Whittam, T.S.; Sawyer, S.A.; Musser, M.L. Microbiol. 16, 509-519, 1995
A:Title: Molecular population genetic analysis of the streptokinase gene of Streptococcus pyogenes
A:Reference number: S77671; MUID:96037795; PMID:7565111
A:Accession: S77680
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-128 <KAP>
A:Cross-references: UNIPROT:Q54687; UNIPARC:UPI00000BD2D3; ENML:U25862; NID:g818926; PID:g818926
A:Experimental source: strain ET51/M17

C;Genetics:

A;Gene: ska

C;Superfamily: streptokinase

C;Keywords: hydrolase; plasminogen activator; virulence

Alignment Scores:

Pred. No.: 3.86e-31 Length: 128
Score: 569.00 Matches: 110
Percent Similarity: 91.41% Conservative: 7
Best Local Similarity: 85.94% Mismatches: 11
Query Match: 20.95% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-9 (1-1541) x S77680 (1-128)

QY 453 GTGCGGTTAGACCATATAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 512

Db 1 ValArgValArgProTyrLysGluLysProIleGlnThrProAlaLysSerValAspVal 20

QY 513 GAATATACCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTCAGACAGGTCTCAAA 572

Db 21 ArgTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProValLeuLys 40

QY 573 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTACTA 632

Db 41 AsnThrLysLeuLysThrLeuAlaIleGlyGlyThrValThrSerGlnGluLeuLeu 60

QY 633 GCTCAAGCACAAGCATTTTAAACAAAACCAACCCAGGCTATACGATTTTATGACCGTAC 692

Db 61 AlaGlnAlaGlnSerIleLeuAsnGluSerHisProAspTyrThrIleTyrGluArgAsp 80

QY 693 TCCTCAATCGTCACATGACATGACATTTCCGTACGATTTTACCAATGGATCAAGAG 752

Db 81 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 100

QY 753 TTTACTTACCGTGTAAAAATCGGAAACCAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 812

Db 101 PheThrTyrHisIleLysAspArgGluGlnAlaTyrGlyLeuAsnLysLysSerGlyGln 120

QY 813 AATGAAGAAATAAACACACTGAC 836

Db 121 GluGluLysThrAsnAsnThrAsp 128

RESULT 13

S77688

streptokinase A (EC 3.4.-.-) (allele 19) - Streptococcus pyogenes (fragment)

C;Species: Streptococcus pyogenes

C;Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004

C;Accession: S77688

R;Kaput, V.; Kanjilal, S.; Hamrick, M.R.; Li, L.L.; Whittam, T.S.; Sawyer, S.A.; Musser, M.L. Microbiol. 16, 509-519, 1995

A;Title: Molecular population genetic analysis of the streptokinase gene of Streptococcus

A;Reference number: S77671; MUID:96037795; PMID:7565111

A;Accession: S77688

A;Status: translation not shown

A;Molecule type: DNA

A;Residues: 1-128 <KAP>

A;Cross-references: UNIPROT:Q54695; UNIPARC:UPI00000BD8DC; EMBL:U25870; NID:g818942; PII

A;Experimental source: strain ET76/M72

C;Genetics:

A;Gene: ska

C;Superfamily: streptokinase

C;Keywords: hydrolase; plasminogen activator; virulence

Alignment Scores:

Pred. No.: 2.53e-30 Length: 128
Score: 557.00 Matches: 107
Percent Similarity: 89.84% Conservative: 8
Best Local Similarity: 83.59% Mismatches: 13
Query Match: 20.51% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-9 (1-1541) x S77688 (1-128)

QY 453 GTGCGGTTAGACCATATAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 512

Db 1 ValArgValArgProTyrLysGluLysProIleGlnThrProAlaLysSerValAspVal 20

QY 513 GAATATACCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTCAGACAGGTCTCAAA 572

Db 21 ArgTyrThrValGlnPheThrProLeuAsnProAspAspPheLysProValLeuLys 40

QY 573 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTACTA 632

Db 41 AspThrLysLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 60

QY 633 GCTCAAGCACAAGCATTTTAAACAAAACCAACCCAGGCTATACGATTTTATGACCGTAC 692

Db 61 AlaGlnAlaGlnSerIleLeuLysGluSerHisProAspTyrThrIleTyrGluArgAsp 80

QY 693 TCCTCAATCGTCACATGACATGACATTTCCGTACGATTTTACCAATGGATCAAGAG 752

Db 81 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProThrAspGlnGlu 100

QY 753 TTTACTTACCGTGTAAAAATCGGAAACCAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 812

Db 101 PheThrTyrHisValLysAsnArgGluGlnAlaTyrLysAlaAsnSerLysThrAspVal 120

QY 813 AATGAAGAAATAAACACACTGAC 836

Db 121 LysGluLysThrAsnAsnThrAsp 128

RESULT 14

S77679

streptokinase A (EC 3.4.-.-) (allele 10) - Streptococcus pyogenes (fragment)

C;Species: Streptococcus pyogenes

C;Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004

C;Accession: S77679

R;Kaput, V.; Kanjilal, S.; Hamrick, M.R.; Li, L.L.; Whittam, T.S.; Sawyer, S.A.; Musser, M.L. Microbiol. 16, 509-519, 1995

A;Title: Molecular population genetic analysis of the streptokinase gene of Streptococcus

A;Reference number: S77671; MUID:96037795; PMID:7565111

A;Accession: S77679

A;Status: translation not shown

A;Molecule type: DNA

A;Residues: 1-128 <KAP>

A;Cross-references: UNIPROT:Q54686; UNIPARC:UPI00000BD8D6F; EMBL:U25861; NID:g818924; P

A;Experimental source: strain ET50/M43

C;Genetics:

A;Gene: ska

C;Superfamily: streptokinase

C;Keywords: hydrolase; plasminogen activator; virulence

Alignment Scores:

Pred. No.: 5.53e-30 Length: 128
Score: 552.00 Matches: 106
Percent Similarity: 89.06% Conservative: 8
Best Local Similarity: 82.81% Mismatches: 14
Query Match: 20.32% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-9 (1-1541) x S77679 (1-128)

QY 453 GTGCGGTTAGACCATATAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 512

Db 1 ValArgValArgProTyrLysGluLysProIleGlnThrProAlaLysSerValAspVal 20

QY 513 GAATATACCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTCAGACAGGTCTCAAA 572

Db 21 ArgTyrAlaValGlnPheThrProLeuAsnProAspAspPheThrProValLeuLys 40

QY 573 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTACTA 632

Db 41 AspThrLysLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 60

QY 633 GCTCAAGCACAAGCATTTTAAACAAAACCAACCCAGGCTATACGATTTTATGACCGTAC 692

```

Db      61 AlaGlnAlaGlnSerIleLeuIleGluSerHisProAsnTyrThrIleHisGluArgAsp 80
      693 TCCTCAATCGTCACATGACATGACATTTTCCTGACGATTTTACCAATGATGATCAAGAG 752
      81 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProThrAspGlnGlu 100
      753 TTTACTTACCGTGTAAAAATCGGAACAACAGCTTATAGGATCAATAAAAAATCTGCTCG 812
      101 PheThrTyrHisValLysAsnArgGluGlnAlaTyrLysAlaAsnSerLysThrGlyIle 120
      813 AATGAAGAAATAAACCACTGAC 836
      121 LysGluLysThrAsnAsnThrAsp 128

RESULT 15
A43908
Fibronectin - African clawed frog
C/Species: Xenopus laevis (African clawed frog)
C/Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 12-Feb-1999
C/Accession: A43908
R/DeSimone, D.W.; Norton, P.A.; Hynes, R.O.
Dev. Biol. 149, 357-369, 1992
A/Title: Identification and characterization of alternatively spliced fibronectin mRNAs
A/Reference number: A43908; MUID:92111942; PMID:1730390
A/Accession: A43908
A/Status: nucleic acid sequence not shown; not compared with conceptual translation
A/Molecule type: mRNA
A/Residues: 1-2481 <DES>
A/Cross-references: UNIPARC:UPI0000177AE7; GB:M77820
C/Note: sequence extracted from NCBI backbone (NCBIP:77473)
A/Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II repe
C/Keywords: duplication; extracellular matrix; glycoprotein; heterodimer
F/58-90/Domain: fibronectin type I repeat homology <1F1>
F/100-138/Domain: fibronectin type I repeat homology <1F2>
F/144-182/Domain: fibronectin type I repeat homology <1F3>
F/189-228/Domain: fibronectin type I repeat homology <1F4>
F/234-273/Domain: fibronectin type I repeat homology <1F5>
F/309-343/Domain: fibronectin type I repeat homology <1F6>
F/361-402/Domain: fibronectin type II repeat homology <2F1>
F/421-462/Domain: fibronectin type II repeat homology <2F2>
F/471-509/Domain: fibronectin type I repeat homology <1F7>
F/519-556/Domain: fibronectin type I repeat homology <1F8>
F/562-600/Domain: fibronectin type I repeat homology <1F9>
F/610-693/Domain: fibronectin type III repeat homology <FN3A>
F/719-801/Domain: fibronectin type III repeat homology <FN3B>
F/810-891/Domain: fibronectin type III repeat homology <FN3C>
F/906-988/Domain: fibronectin type III repeat homology <FN3D>
F/996-1077/Domain: fibronectin type III repeat homology <FN3E>
F/1086-1165/Domain: fibronectin type III repeat homology <FN3F>
F/1173-1258/Domain: fibronectin type III repeat homology <FN3G>
F/1266-1349/Domain: fibronectin type III repeat homology <FN3H>
F/1357-1440/Domain: fibronectin type III repeat homology <FN3I>
F/1448-1530/Domain: fibronectin type III repeat homology <FN3J>
F/1538-1620/Domain: fibronectin type III repeat homology <FN3K>
F/1615-1617/Region: cell attachment (R-G-D) motif
F/1632-1714/Domain: fibronectin type III repeat homology <FN3L>
F/1722-1804/Domain: fibronectin type III repeat homology <FN3M>
F/1812-1894/Domain: fibronectin type III repeat homology <FN3N>
F/1904-1985/Domain: fibronectin type III repeat homology <FN3O>
F/1993-2075/Domain: fibronectin type III repeat homology <FN3P>
F/2197-2277/Domain: fibronectin type III repeat homology <FN3Q>
F/2301-2340/Domain: fibronectin type I repeat homology <1F10>
F/2346-2383/Domain: fibronectin type I repeat homology <1F11>
F/2390-2425/Domain: fibronectin type I repeat homology <1F12>
F/55-81,79-90,100-128,126-138,144-172,170-182,189-218,216-228,234-263,261-273,309-336,33
2373,2371-2383,2390-2416,2414-2425/Disulfide bonds: #status predicted
F/2459/Disulfide bonds: interchain (to 2463) #status predicted
F/2463/Disulfide bonds: interchain (to 2459) #status predicted

```

Alignment Scores:

Pred. No.:	2,53e-29	Length:	2481
Score:	544.00	Matches:	95

```

Percent Similarity: 81.75% Conservative: 17
Best Local Similarity: 69.34% Mismatches: 19
Query Match: 20.03% Indels: 6
DB: 2 Gaps: 3
US-09-940-235-9 (1-1541) x A43908 (1-2481)
QY 1143 AAGCGACCCGAGAGGAGAGATGCTAGCTACCATTTA-----GCTGGTGGTGGC 1190
Db 158 ArgArgProHis---GluThrGlyGlyTyrMetLeuGluCysValCysLeuGlyAsnGly 176
QY 1191 CAGCGCCAA---CAGATTGTACCCATAGCTGAGAGAGTGTGTTTGTATCATGCTGCTGGGACT 1247
Db 177 LysGlyGluTrpThrCysLysProValAlaGluArgCysTyrAspAsnThrAlaGlyThr 196
QY 1248 TCCTATGCTGCTCGGAGAAACGTCGGAGAGCCCTACCAAGCTGGATGGTAGATTGT 1307
Db 197 SerTyrValValGlyGlnThrTrpGluLysProTyrGlnGlyTrpMetValAspCys 216
QY 1308 ACTTGCTGCGGAGAGCGAGCGACCATTCGACTTCTAGAAATAGATGCAACGAT 1367
Db 217 ThrCysLeuGlyGluGlyAsnGlyArgIleThrCysSerSerLysAsnArgCysAsnAsp 236
QY 1368 CAGGACACAAGGACATCCTATAGAAATTGGAGACACCTGGAGCAAGAGGATAATCGAGGA 1427
Db 237 GlnAspThrLysThrSerTyrArgIleGlyAspThrTrpSerLysThrAspThrArgGly 256
QY 1428 AACCTGCTCCAGTGCATCTGCGACAGCAACGGCCGAGAGAGTGGAGTGTGTAGAGGCAC 1487
Db 257 AsnLeuLeuGlnCysIleCysThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHis 276
QY 1488 ACCTCTGTGTCAGACACATCGAGCGGATCTGGCCCTTCACCGATGTTCT 1538
Db 277 SerSerAlaGlnAlaThrGlyThrGlySerAsnProIleThrAsnIleGln 293

```

Search completed: January 28, 2006, 02:37:22

Job time : 38.6156 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 28, 2006, 01:48:09 ; Search time 166.071 Seconds
(without alignments)
13093.422 Million cell updates/sec

Title: US-09-940-235-9
Perfect score: 2716
Sequence: 1 ttgtttaactttaagaagg.....ccttcaccgatgttcgtag 1541

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 4332886

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US0940235/runat_27012006_144218_27563/app_query.fasta_1.7708
-DB=UniProt -QFMT=fastan -SUFFIX=rup -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=plo -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US0940235 @CGN_1_1_1359 @runat_27012006_144218_27563 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : UniProt_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1984	73.0	440	1 STRP_STRPQ	P00779 streptococc
2	1955	72.0	414	2 Q53284 STRPQ	Q53284 streptococc
3	1943	71.5	436	2 Q6UK57 STRPQ	Q6UK57 streptococc
4	1942	71.5	440	1 STRP_STRS1	P10519 streptococc
5	1814	66.8	440	2 Q7X0Y1 STRPY	Q7X0Y1 streptococc
6	1802	66.3	440	2 Q8X5R8 STRP3	Q8X5R8 streptococc
7	1802	66.3	440	1 STRP_STRPY	P10520 streptococc
8	1794	66.1	440	2 Q532X6 STRPY	Q532X6 streptococc
9	1794	66.1	440	2 Q7X0Y2 STRPY	Q7X0Y2 streptococc
10	1787	65.8	440	2 Q7X0Y7 STRPY	Q7X0Y7 streptococc
11	1758	64.7	440	2 Q5X9T6 STRP6	Q5X9T6 streptococc
12	1758	64.6	440	2 Q8NZA6 STRP8	Q8NZA6 streptococc
13	1755	64.4	440	2 Q7X0Y3 STRPY	Q7X0Y3 streptococc
14	1723	63.4	440	2 Q7X0Y0 STRPY	Q7X0Y0 streptococc
15	1715	63.1	440	2 Q7X0Y5 STRPY	Q7X0Y5 streptococc
16	1715	63.1	440	2 Q7X0Y5 STRPY	Q7X0Y5 streptococc

17	1708	62.9	440	2	Q7X0X8 STRPY	Q7X0X8 streptococc
18	1686	62.1	440	2	Q7X0X9 STRPY	Q7X0X9 streptococc
19	1677	61.7	440	2	Q7X0X7 STRPY	Q7X0X7 streptococc
20	1670	61.5	440	2	Q57391 STRPY	Q57391 streptococc
21	1657	61.0	440	2	Q7X0Y6 STRPY	Q7X0Y6 streptococc
22	1652	60.8	440	1	STRQ_STRPY	P96471 streptococc
23	1624	59.8	432	2	Q7X0Y4 STRPY	Q7X0Y4 streptococc
24	728	26.8	141	2	Q7X0X2 STRPQ	Q7X0X2 streptococc
25	727	26.8	141	2	Q7X0X3 STRPQ	Q7X0X3 streptococc
26	724	26.7	141	2	Q7X0W1 STRPQ	Q7X0W1 streptococc
27	723	26.6	141	2	Q7WS87 STRPQ	Q7WS87 streptococc
28	717	26.4	141	2	Q7X0W3 STRPQ	Q7X0W3 streptococc
29	717	26.4	141	2	Q7X0X4 STRPQ	Q7X0X4 streptococc
30	715	26.3	141	2	Q7X0W5 STRPQ	Q7X0W5 streptococc
31	714	26.3	141	2	Q7X0X5 STRPQ	Q7X0X5 streptococc
32	712	26.2	141	2	Q7X0X6 STRPQ	Q7X0X6 streptococc
33	711	26.2	141	2	Q7X0W2 STRPQ	Q7X0W2 streptococc
34	711	26.2	141	2	Q7X0W9 STRPQ	Q7X0W9 streptococc
35	710	26.1	141	2	Q7X0W6 STRPQ	Q7X0W6 streptococc
36	707	26.0	141	2	Q7X0W0 STRPQ	Q7X0W0 streptococc
37	705	26.0	141	2	Q7X0V5 STRPY	Q7X0V5 streptococc
38	704	25.9	141	2	Q7X0X0 STRPQ	Q7X0X0 streptococc
39	701	25.8	141	2	Q7X0W7 STRPQ	Q7X0W7 streptococc
40	697	25.7	141	2	Q7X0W4 STRPQ	Q7X0W4 streptococc
41	696	25.6	141	2	Q7X0W8 STRPQ	Q7X0W8 streptococc
42	693	25.5	141	2	Q7X0X1 STRPQ	Q7X0X1 streptococc
43	690	25.4	141	2	Q7X0R2 STRPY	Q7X0R2 streptococc
44	686	25.3	141	2	Q7X0S4 STRPY	Q7X0S4 streptococc
45	684	25.2	141	2	Q7X0S0 STRPY	Q7X0S0 streptococc

ALIGNMENTS

RESULT 1

STRP_STRPQ

ID STRP_STRPQ STANDARD; PRT; 440 AA.

AC P00779;

DT 21-JUL-1986 (Rel. 01, Created)

DT 01-APR-1988 (Rel. 07, Last sequence update)

DT 10-MAY-2005 (Rel. 47, Last annotation update)

DE Streptokinase C precursor.

GN Name=skc;

OS Streptococcus equisimilis.

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

OC Streptococcus.

OX NCBI_TaxID=119602;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=H46A;

RX MEDLINE=85232082; Pubmed=2989113; DOI=10.1016/0378-1119(85)90145-3;

RA Malke H., Roe B.A., Ferretti J.J.;

RT "Nucleotide sequence of the streptokinase gene from Streptococcus

equisimilis H46A.";

RL Gene 34:357-362(1985).

RN [2]

RP PROTEIN SEQUENCE OF 27-440.

RX MEDLINE=83127125; Pubmed=6760891;

RA Jackson K.W., Tang J.;

RT "Complete amino acid sequence of streptokinase and its homology with

serine proteases.";

RL Biochemistry 21:6620-6625(1982).

CC -/- FUNCTION: This protein is not a protease, but it activates

CC plasminogen by complexing with it. As a potential virulence

CC factor, it is thought to prevent the formation of effective fibrin

CC barriers around the site of infection, thereby contributing to the

CC invasiveness of the cells.

CC -----

CC This Swiss-Prot entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use as long as its content is in no way modified and this statement is not

CC removed.

CC -----

DR EMBL; K02986; AAA26974.1; -; Genomic DNA.
 DR EMBL; X72832; CAA51351.1; -; Genomic DNA.
 DR PIR; A00967; BZSO.
 DR PIR; A22801; A22801.
 DR PDB; 1BML; X-ray; C/D=38-399.
 DR PDB; 1L4D; X-ray; B=40-173.
 DR PDB; 1L4Z; X-ray; B=27-173.
 DR PDB; 1LQZ; X-ray; A/B/C/D=177-314.
 DR SMC; P00779; 38-398.
 DR InterPro; IPR004093; Staphylokinase.
 DR InterPro; IPR008124; Streptokinase.
 DR Pfam; PF02821; Staphylokinase; 3.
 DR PRINTS; PR01753; STREPKINASE.
 DR 3D-structure; Direct protein sequencing; Plasminogen activation;
 KW Signal; Virulence.
 FT SIGNAL 1 26
 FT CHAIN 27 440 Streptokinase C.
 FT VARIANT 195 195 L -> D.
 FT VARIANT 207 207 D -> L.
 FT CONFLICT 298 300 EKY -> LEYK (in Ref. 2).
 FT CONFLICT 438 438 N -> D (in Ref. 2).
 FT STRAND 180 180
 FT STRAND 184 194
 FT TURN 199 200
 FT TURN 203 204
 FT STRAND 205 205
 FT STRAND 209 214
 FT TURN 216 217
 FT STRAND 219 221
 FT HELIX 222 236
 FT TURN 238 239
 FT STRAND 240 252
 FT TURN 253 254
 FT STRAND 259 260
 FT STRAND 268 270
 FT STRAND 275 275
 FT STRAND 278 280
 FT TURN 282 284
 FT STRAND 287 289
 FT STRAND 292 304
 FT TURN 305 306
 SQ SEQUENCE 440 AA; 50140 MW; 8PCLF22648ACC77A CRC64;
 Alignment Scores:
 Pred. No.: 2.21e-133 Length: 440
 Score: 1984.00 Matches: 383
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 73.05% Indels: 0
 DB: 1 Gaps: 0
 US-09-940-235-9 (1-1541) x STRP_STREQ (1-440)
 QY 33 ATAGCTGCTCGAATGGCTACTAGATCGTCTCTGTAATAACAGACCAATTTGGTTGT 92
 DB 27 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal 46
 QY 93 ACCGTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAATC 152
 DB 47 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 66
 QY 153 GATCTAACATCAGCACCTGCTCATGGAGGAAGAGACAGCAAGCGTCTAAGTCAAAATCA 212
 DB 67 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 86
 QY 213 AAACCATTTGCTACTGATAGTGGCCGATGTCACATAACTTGAGAAAGCTGACTACTA 272
 DB 87 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 106
 QY 273 AAGGCTATTCAAGAACAAATTGATCCGTAAACGTCCACAGTAACGACGACTACTTTGAGGTC 332
 DB 107 LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 126

QY 333 ATTGATTTTGAAGCGATGCAACATTACTGATCGAAACGCGAGGTCTACTTTGCTGAC 392
 DB 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146
 QY 393 AAAGATGGTTCGGTAACTTGGCGACCAACCTGTCCAAGAAATTTTGTCTAAGCGGACAT 452
 DB 147 LysAspGlySerValThrLeuProThrGlnProValGlnGlnGluPheLeuLeuSerGlyHis 166
 QY 453 GTGGCGTTAGACCATATAAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 512
 DB 167 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 186
 QY 513 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGTCTCAAA 572
 DB 187 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 206
 QY 573 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGAATTACTA 632
 DB 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 226
 QY 633 GCTCAAGCAAAAGCATTTTAAACAAAAACCCAGGCTATACGATTTATGAAACGTGAC 692
 DB 227 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 246
 QY 693 TCCTCAATCGTCACTCATGACATGACATTTTCGTCAGCATTTTACCAATGGATCAAGAG 752
 DB 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
 QY 753 TTTACTTACCCTGTTAAAAATCGGAAACCAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 812
 DB 267 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 286
 QY 813 AATGAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTAATCGTCTCTTAAAAAGGG 872
 DB 287 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 306
 QY 873 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCAACCAATACGTT 932
 DB 307 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326
 QY 933 GATGTCGATACCAACGAATTGCTTAAAGTGACAGCTCTTAACAGCTAGCGAACGTAAC 992
 DB 327 AspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
 QY 993 TTAGACTTCAGAGATTATACGATCTCTGATGTAAGCTAAACCTCTTACAAACAATCTC 1052
 DB 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
 QY 1053 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAAGTAGAGGATATACGATGAC 1112
 DB 367 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 386
 QY 1113 ACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAGGAGAGAGATGCTAGCTAC 1172
 DB 387 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 406
 QY 1173 CATTTAGCT 1181
 DB 407 HisLeuAla 409
 RESULT 2
 Q53284 STREQ
 ID Q53284 STREQ PRELIMINARY; PRT; 414 AA.
 AC Q53284;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE SKC-2.
 OS Streptococcus equisimilis.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=119602;
 RN [1]

TGGTT 92

Db 23 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal 42
Qy 93 AGCGTTGCTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC 152
Db 43 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPheGluIle 62
Qy 153 GATCTAACATCACGACCTGCTCATCGAGGAAAGACAGACAAAGCTTAAAGTCCMAATCA 212
Db 63 AspLeuThrSerGlnProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 82
Qy 213 AAACATTGCTGATGATGGCGGATGTCAATAACTTGAGAAAGCTGACTTACTPA 272
Db 83 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu 102
Qy 273 AAGGCTATTCAAGACAAATTGATCCCTAACGTCACGTCACGACGACTTGTAGGTC 332
Db 103 LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 122
Qy 333 ATTGATTTCGAAGCGATGCAACCAATTACTGATCGAAACGCAAGTCTACTTGTGAC 392
Db 123 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlnLysValTyrPheAlaAsp 142
Qy 393 AAAGATGTTGCTGTAACCTTCGCGACCACTGTCGAAGAATTTTGTAAAGCGGACAT 452
Db 143 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuLysGlyHis 162
Qy 453 GTGCGGCTAGACCATATAAGAAAAACCAATACAAACCAAGCGAAATCTGTGATGTG 512
Db 163 ValArgValArgProTyrLysGluLysProValGlnAsnGlnAlaLysSerValAspVal 182
Qy 513 GAATATCTGTCAGTTCCTCCCTTAAACCTCGATGACGATTCAGACGAGTCTCAAA 572
Db 183 GluTyrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuLys 202
Qy 573 GATACTAAGCTATTGAAACACTAGCTATCGGTGACCATCACATCTCAAGAATTACTPA 632
Db 203 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 222
Qy 633 GCTCAAGCACAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTATGAAGTGAC 692
Db 223 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp 242
Qy 693 TCCTCAATCGTCACCTCATGACAATGACATTTTCGTGACGATTTTACCAATGATCAAGAG 752
Db 243 SerSerIleValThrHisAspLysAspIlePheArgThrIleLeuProMetAspGlnGlu 262
Qy 753 TTTACTTACCGTGTAAAAATCGGAAACAAGCTTATAGGATCAATAAAAAATCTGCTGTG 812
Db 263 PheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsnLysLysSerGlyLeu 282
Qy 813 AATGAAGAATAAACAACACTGACCTGATCTCTGAGAAATATATACGTCCTTAAAAAGGG 872
Db 283 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 302
Qy 873 GAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCCACCATCAAAATCGTT 932
Db 303 GluLysProTyrAspProPheAspArgSerHisLysLysLeuPheThrIleLysTyrVal 322
Qy 933 GATCTGATCAACAAGAAATTGCTAAAAAGTGAAGAGCTCTTAAACAGCTAGCCAAAGTAAAC 992
Db 323 AspValAsnThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 342
Qy 993 TTAGACTTCAGAGATTTATAGCTCTCGTGATAAAGGCTTAAACTACTCTACCAACATCTC 1052
Db 343 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 362
Qy 1053 GATGCTTTGCTATATGCTATACCTTAACTGGAAGAGTGAAGATAATCACCATGAC 1112
Db 363 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 382
Qy 1113 ACCAACCGGTATCAATACCGTTTATATGGCGAAGCGACCGGAGAGAGATCTCTAGCTAC 1172

Db 383 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 402
Qy 1173 CATTTAGCT 1181
Db 403 HisLeuAla 405
RESULT 4
STRP_STRS1
ID STRP_STRS1 STANDARD; PRT; 440 AA.
AC P10519;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Streptokinase G precursor.
GN Name=skg;
OS Streptococcus sp. (strain 19909).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=69017;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RX MEDLINE=89160285; PubMed=2922269;
RA Walter F., Siegel M., Malke H.;
RT "Nucleotide sequence of the streptokinase gene from a group-G
RT Streptococcus";
RL Nucleic Acids Res. 17:1262-1262(1989).
CC -!- FUNCTION: This protein is not a protease, but it activates
CC plasminogen by complexing with it. As a potential virulence
CC factor, it is thought to prevent the formation of effective fibrin
CC barriers around the site of infection, thereby contributing to the
CC invasiveness of the cells.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; X13400; CAA31766.1; -; Genomic_DNA.
DR PIR; S02723; S02723.
DR HSP; P00779; 1QOR.
DR SMR; P10519; 63-398.
DR InterPro; IPR004093; Staphylokinase.
DR InterPro; IPR008124; Streptokinase.
DR Pfam; PF02821; Staphylokinase; 3.
DR PRINTS; PR01753; STREPKINASE.
KW Plasminogen activation; Signal; Virulence.
FT SIGNAL 1 26
FT CHAIN 27 440 Streptokinase G.
SQ SEQUENCE 440 AA; 50199 MW; 5521P8825FE1B6EA CRC64;
Alignment Scores:
Pred. No.: 2,24e-130 Length: 440
Score: 1942.00 Matches: 375
Percent Similarity: 98.43% Conservative: 2
Best Local Similarity: 97.91% Mismatches: 6
Query Match: 71.50% Indels: 0
DB: Gaps: 1
US-09-940-235-9 (1-1541) x STRP_STRS1 (1-440)
Qy 33 ATAGCTGGTCCTGAATGGCTACTAGATCGTCTTCTGTAATAACAGCCAAATGGTTGTT 92
Db 27 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal 46
Qy 93 AGCGTTGCTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC 152
Db 47 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPheGluIle 66
Qy 153 GATCTAACATCACGACCTGCTCATCGAGGAAAGACAGACAAAGCTTAAAGTCCMAATCA 212
Db 67 AspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 86

```
QY 213 AAACCAATTTGCTACTGATAGTGGCGGCGATGTCAATAAATCTTGAGAAAGCTGACTTACTA 272
Db 87 LysLeuPheAlaThrAspSerGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu 106
QY 273 AAGGCTATTCAAGCAAAATGATCGCTAAAGTCCACAGTAACGACGACTACTTTGAGGTC 332
Db 107 LysAlaIleGlnGluLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 126
QY 333 ATTGATTTTGCAGCGATGCAACCATTTACTGATCGAAACCGCAAGGCTCTACTTTGCTGAC 392
Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146
QY 393 AAGATGTTTGGTAACTTTGCCGACCCCAACCTGTCCCAAGATTTTTCGTAAGCGGACAT 452
Db 147 LysAspGlySerValThrLeuProIleGlnProValGlnPheLeuLeuLysGlyHis 166
QY 453 GTGCGGTTAGACCATATAAAGAAACCAATACAAACCAAGGAAATCTGTTGATGTG 512
Db 167 ValArgValArgProTyrLysGluLysProValGlnAsnGlnAlaLysSerValAspVal 186
QY 513 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTCAAA 572
Db 187 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProAlaLeuLys 206
QY 573 GATATCAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTACTA 632
Db 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 226
QY 633 GCTCAAGCAAAAGCAATTTTAAACAAACCCAGGCTATACGATTTATGACGTTGAC 692
Db 227 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 246
QY 693 TCTCTCAATCGTCACTCATGACATGACATTTTCGTCAGTATTTTACCAATGGATCAGAG 752
Db 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
QY 753 TTTACTTACCGTGTAAAAATCGGGAACAGCTTTATAGGATCAATAAAAAATCTGGTCTG 812
Db 267 PheThrTyrHisValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 286
QY 813 AATGAAGAAATAAACCACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 872
Db 287 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 306
QY 873 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCCACCATCAATAGCTT 932
Db 307 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326
QY 933 GATGTCGATACCAACCAATTTGCTAAAAAGTGACGAGCTCTTAACAGCTAGCGAACGTAAC 992
Db 327 AspValAsnThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
QY 993 TTGAGCTTCAGAGATTTATACGATCTCTGATGATAGGCTAAATCTCTACACAAATCTC 1052
Db 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
QY 1053 GATGCTTTTGGTATTATGACTATACCTTAATCTGGAAGAGTAGAGGATATCATCGATGAC 1112
Db 367 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 386
QY 1113 ACCAACCGTATCATACCGTTTATATGGCGACGACCGACCGAGGAGGATGCTAGCTAC 1172
Db 397 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 406
QY 1173 CATTTAGCT 1181
Db 407 HisLeuAla 409
RESULT 5
Q7X0Y1 STRPY
ID Q7X0Y1 STRPY PRELIMINARY, PRT; 440 AA.
AC Q7X0Y1,
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
```

```
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DB Streptokinase.
GN Name=ska;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=D488;
RX PubMed=14679231; DOI=10.1128/JB.186.1.110-121.2004;
RA Kalia A., Bessen D.R.;
RT "Natural selection and evolution of streptococcal virulence genes
involved in tissue-specific adaptations.";
RL J. Bacteriol. 186:110-121(2004).
DR EMBL; AY234137; AAP39957.1; -; Genomic_DNA.
DR HSSP; Q53284; 1C4P.
DR SMR; Q7X0Y1; 63-398.
DR GO; GO:0016301; P-kinase activity; IEA.
DR GO; GO:0008243; P-plasminogen activator activity; IEA.
DR GO; GO:0005515; F-protein binding; IEA.
DR InterPro; IPR004093; Staphylokinase.
DR InterPro; IPR008124; Streptokinase.
DR Pfam; PF02821; Staphylokinase; 3.
DR PRINTS; PR01753; STREPKINASE.
KW Kinase.
SQ SEQUENCE 440 AA; 49903 MW; CC4B6B9647043BAC CRC64;

Alignment Scores:
Pred. No.: 3.25e-121 Length: 440
Score: 1814.00 Matches: 347
Percent Similarity: 95.04% Conservative: 17
Best Local Similarity: 90.60% Mismatches: 19
Query Match: 66.79% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-9 (1-1541) x Q7X0Y1_STRPY (1-440)
QY 33 ATAGCTGGTCTGTAATCGCTACTAGATCGTCTTCTGTATAATAACAGCAATGTTGTTT 92
Db 27 IleAlaGlyTyrGlyTyrLeuProAspArgProIleAsnAsnSerGlnLeuValVal 46
QY 93 AGCGTTCCTGTTACTGTTGAGGGACGAATCAAGACATTAGCTTTAAATTTTTTGAATC 152
Db 47 SerMetAlaGlyIleValGluGlyThrAspLysLysValPheIleAsnPheGluIle 66
QY 153 GATCTACATCAGCACCTGCTCATGGAGGAAAGACAGCAAGGCTTAAGTCCAAATCA 212
Db 67 AspLeuThrSerGlnProAlaHisGlyLysThrGluGlnGlyLeuSerProLysSer 86
QY 213 AAACCAATTTGCTACTGATAGTGGCGGCGATGTCAATAAATCTTGAGAAAGCTGACTTACTA 272
Db 87 LysProPheAlaThrAspAsnGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu 106
QY 273 AAGGCTATTCAAGCAAAATGATCGCTTAACGTCCACAGTAACGACGACTACTTTGAGGTC 332
Db 107 LysAlaIleGlnLysGlnLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal 126
QY 333 ATTGATTTTGCAGCGATGCAACCATTTACTGATCGAAACCGCAAGGCTCTACTTTGCTGAC 392
Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146
QY 393 AAGATGTTTGGTAACTTTGCCGACCCCAACCTGTCCCAAGATTTTTCGTAAGCGGACAT 452
Db 147 LysAspGlySerValThrLeuProIleGlnProValGlnPheLeuLeuLysGlyHis 166
QY 453 GTGCGGTTAGACCATATAAAGAAACCAATACAAACCAAGGAAATCTGTTGATGTG 512
Db 167 ValArgValArgProTyrLysGluLysProValGlnAsnGlnAlaLysSerValAspVal 186
QY 513 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTCAAA 572
:::|
```

```
Db 187 LysThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 206
QY 573 GATACAAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAACTACTA 632
Db 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 226
QY 633 GCTCAAGCACAAAGCATTTTAAACAAACCCAGGCTATACGATTTATGATCAACGTGAC 692
Db 227 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyThrIleThrLysGluArgAsp 246
QY 693 TCCTCAATCGTCACATGACATGACATTTCCGTACGATTTTACCAATGATCAAGAG 752
Db 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
QY 753 TTTACTTACCGTGTAAATAATCGGAAACAAGCTTATAGGATCAATAAAAAATCTGCTCG 812
Db 267 PheThrThrArgValLysAspArgGluGlnAlaIleGlyLeuAsnLysSerGlyLeu 286
QY 813 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGCTCTTAAAAAGGG 872
Db 287 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysThrIleLeuLysLysGly 306
QY 873 GAAAGCCGTATGATCCCTTTGATCGCAGTCTCTGAAACTGTTCAACATCAAAATACGTT 932
Db 307 GluSerProThrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysThrVal 326
QY 933 GATGTCGATACCAACGAATTGCTAAAAGTGGAGCAGCTCTTAACAGCTAGCGCAACGTAAC 992
Db 327 AspValAsnThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
QY 993 TTAGACTTTCAGAGATTATACGATCTCTGATAGGCTTAAACTACTCTACCAACAATCTC 1052
Db 347 LeuAspPheArgAspLeuThrAspProArgAspLysAlaLysLeuLeuThrAsnAsnLeu 366
QY 1053 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAGAGTAGAGGATTAATCACGATGAC 1112
Db 367 AspAlaPheAspIleMetAspThrLeuThrGlyLysValGluAspAsnHisAspLys 386
QY 1113 ACCACCGTATCATACCGTTTATATGGCAGCGACCGACCGAGGAGAGATGCTAGCTAC 1172
Db 387 AsnAsnArgIleValThrValThrMetGlyLysArgProLysGlyAlaLysGlySerThr 406
QY 1173 CATTTAGCT 1181
Db 407 HisLeuAla 409
```

RESULT 6

```
Q7X0Y8 STRPY PRELIMINARY; PRT; 440 AA.
AC Q7X0Y8;
DT 01-OCT-2003 (TRENBLrel. 25, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Streptokinase.
GN Name=ska;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=88-019;
RA PubMed=14679231; DOI=10.1128/JB.186.1.110-121.2004;
RX Kalia A., Bessen D.E.;
RT "Natural selection and evolution of streptococcal virulence genes
involved in tissue-specific adaptations.";
RL J. Bacteriol. 186:110-121(2004).
DR EMBL; AY234129; AAP39949.1; -; Genomic_DNA.
DR HSPF; Q53284; IC4P.
DR SMR; Q7X0Y8; 63-398.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008243; F:plasminogen activator activity; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
```

```
DR InterPro; IPR004093; Staphylokinase.
DR InterPro; IPR008124; Streptokinase.
DR Pfam; PF02821; Staphylokinase; 3.
DR PRINTS; PR01753; STREPKINASE.
KW Kinase.
SQ SEQUENCE 440 AA; 49885 MW; 69DB44F4026E3975 CRC64;
```

Alignment Scores:

```
Pred. No.: 2 34e-120 Length: 440
Score: 1802.00 Matches: 347
Percent Similarity: 94.52% Conservative: 15
Best Local Similarity: 90.60% Mismatches: 21
Query Match: 66.35% Indels: 0
DB: 2 Gaps: 0
```

US-09-940-235-9 (1-1541) x Q7X0Y8_STRPY (1-440)

```
QY 33 ATAGCTGGTCTCGAATGGCTACTAGATCGCTCTTCTGTAAATAACACCGCAATTTGGTTGTT 92
Db 27 IleAlaGlyThrGlyTrpLeuProAspArgProValAsnAsnSerGlnLeuVal 46
QY 93 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC 152
Db 47 SerMetAlaGlyIleValGluGlyThrAspLysLysValPheIleAsnPheGluIle 66
QY 153 GATCTAATCATCACGACCTGCTCATGAGGAAAGACAGCAAGCGCTTAAAGTCCAAATCA 212
Db 67 AspLeuThrSerGlnHisAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 86
QY 213 AAACCAATTTGCTACTGATAGTGGCGGATGTCACATAAACTTGAGAAAGCTGACTACTA 272
Db 87 LysProPheAlaThrAspAsnGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu 106
QY 273 AGGCTATTCAAGAACAAATTGATCGCTTAAGTCCACAGTAACGACGACTACTTTGAGGTC 332
Db 107 LysAlaIleGlnLysGlnLeuIleAlaAsnValHisSerAsnAspGlyThrPheGluVal 126
QY 333 ATTGATTTTGCAGCGCATGCCACCATTAAGTCGAAACGGCAAGGTCTACTTTGCTGAC 392
Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValThrPheAlaAsp 146
QY 393 AAAGATGTTTGGTAACTTTCCTTAAACCTTGCACCACTGTCCAAGATTTTGTAAAGCGACAT 452
Db 147 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 166
QY 453 GTGCGGTTAGACCATATAAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 512
Db 167 ValArgValArgProThrLysGluLysProValGlnAsnGlnAlaLysSerValAspVal 186
QY 513 GAATATACGTACAGTTTACTCCCTTAAACCTTGAAGTATGACGATTTTCAGCCAGGCTCAAA 572
Db 187 LysThrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 206
QY 573 GATACTAAGCTATTGAAACACATAGCTATCGGTGACACCATCACATCTCAAGAAATACTA 632
Db 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 226
QY 633 GCTCAAGCACAAAGCATTTTAAACAAACCCAGGCTATACGATTTTATGAACCGTGAC 692
Db 227 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyThrIleThrLysGluArgAsp 246
QY 693 TCCTCAATCGTCATCATGACATGATTTTCGTCAGATTTTACCAATGGATCAAGAG 752
Db 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
QY 753 TTTACTTACCGTGTAAATAATCGGAAACAAGCTTATAGGATCAATAAAAAATCTGCTCG 812
Db 267 PheThrThrArgValLysAspArgGluGlnAlaIleGlyIleAsnLysLysSerGlyLeu 286
QY 813 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGCTCTTAAAAAGGG 872
Db 287 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysThrIleLeuLysLysGly 306
```


Qy	873	GAAAGCGGTATGATCCCTTTGATCGCAGTCACTCTGAAACATGTTCACCATCAAAATACGTT	933
Db	307	GlusErProTyAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTy-Val	326
Qy	933	GATGTGCATACCAACGAATTCCTAAAAGTCAGCAGCTCTTAAACAGCTAGCGACGTAAC	992
Db	327	AspValAnThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn	346
Qy	993	TTAGACTTCAGAGATTTATACGATCTCGTGATAGGCTAAAGCTAACTACTCTACAACAATCTC	1052
Db	347	LeuAspPheArgAspLeuTyAspProCysAspLysAlaLysLeuLeuTyAsnAnLeu	366
Qy	1053	CATGCTTTTGGTATTATGCACTATACCTTAACTGGAAAGATAGAGGATTAATCAGCATGAC	1112
Db	367	AspAlaPheAspIleMetAspTy-ThrLeuThrGlyLysValGluAspAsnHisAspLys	386
Qy	1113	ACCAACGCTATCATACCGTTTATATGGCAAGCGACCGCAAGAGAGAGAAATGCTAGCTAC	1172
Db	387	AsnAnArgIleValThrValTyMetGlyLysArgProLysGlyAlaLysGlySerTy	406
Qy	1173	CATTTAGCT 1181	
Db	407	HisLeuAla 409	
RESULT 7			
QID	Q8KSR8_STRP3	PRELIMINARY; PRT; 440 AA.	
ID	Q8KSR8_STRP3	PRELIMINARY; PRT; 440 AA.	
AC	Q8KSR8; 079W73;		
DT	01-OCT-2002 (TrEMBLrel. 22, Created)		
DT	01-OCT-2002 (TrEMBLrel. 22, Last sequence update)		
DT	13-SEP-2005 (TrEMBLrel. 31, Last annotation update)		
DE	Streptokinase A.		
GN	Name=ska; OrderedLocusNames=SPs1700, SPyM3_1698;		
OS	Streptococcus pyogenes (serotype M3).		
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;		
OC	Streptococcus.		
OX	NCBI_TaxID=301448;		
RN	[1]		
RN	NUCLEOTIDE SEQUENCE.		
RC	STRAIN=MGAS315 / Serotype M3;		
RC	MEDLINE=22133808; PubMed=12122206; DOI=10.1073/pnas.152298499;		
RA	Beres S.B., Sylva G.L., Barbian K.D., Iei B., Hoff J.S.,		
RA	Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,		
RA	Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,		
RA	Schlievert P.M., Mueser J.M.;		
RT	"Genome sequence of a serotype M3 strain of group A Streptococcus:		
RT	phase-encoded toxins, the high-virulence phenotype, and clone		
RT	emergence.";		
RL	Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).		
RN	[2]		
RN	NUCLEOTIDE SEQUENCE.		
RC	STRAIN=SSI-1 / Serotype M3;		
RC	MEDLINE=22683278; PubMed=12799345; DOI=10.1101/gr.1096703;		
RA	Nakagawa I., Kurokawa K., Yamashita A., Nakata M., Tomiyasu Y.,		
RA	Okahashi N., Kawabata S., Yamazaki K., Shiba T., Yasunaga T.,		
RA	Hayashi H., Hattori M., Hamada S.;		
RT	"Genome sequence of an M3 strain of Streptococcus pyogenes reveals a		
RT	large-scale genomic rearrangement in invasive strains and new insights		
RT	into phase evolution.";		
RL	Genome Res. 13:1042-1055(2003).		
RL	EMBL; AE014169; AAM80305.1; -; Genomic_DNA.		
DR	EMBL; BA000334; BAC64795.1; -; Genomic_DNA.		
DR	HSSP; P00779; 1L4D.		
DR	SMR; Q8KSR8; 63-398.		
DR	GO; GO:0016301; F:Kinase activity; IEA.		
DR	GO; GO:0008243; F:Plasminogen activator activity; IEA.		
DR	GO; GO:000515; F:protein binding; IEA.		
DR	InterPro; IPR004093; Staphylokinase.		
DR	InterPro; IPR008124; Streptokinase.		
DR	Pfam; PF02821; Staphylokinase; 3.		
DR	PRINTS; PR01753; STREPKINASE.		
DR	Complete proteome; Kinase.		
SW	SEQUENCE 440 AA; 49911 MW; 7CCCE44F04026E3975 CRC64;		

Alignment Scores:				
Pred. No.:	2,34e-120	Length:	440	
Score:	1802.00	Matches:	347	
Percent Similarity:	94.52%	Conservative:	15	
Best Local Similarity:	90.60%	Mismatches:	21	
Query Match:	66.35%	Indels:	0	
DB:	2	Gaps:	0	
US-09-940-235-9 (1-1541) x Q8K5R8_STRP3 (1-440)				
QY	33	ATAGCTGTGCTCGAATGGCTACTAGACTGCTCTCTGTAAATAACAGCGCAATTGGTTGTT	92	
DB	27	IleAlaGlyTyrGlyTrpLeuProAspArgProValAsnAsnSerGlnLeuValVal	46	
QY	93	AGCGTTGCTGGTACTGTTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAAATC	152	
DB	47	SerMetAlaGlyIleValGluGlyThrAspLysLysValPheIleAsnPheGluIle	66	
QY	153	GATCTAAATCATCAGACCTGCTCATCGAGGAAGAAGACAGACAGCGCTTAAGTCCAAATCA	212	
DB	67	AspLeuThrSerGlnHisAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer	86	
QY	213	AAACCATTTGCTACTGATAGTGGCCCGATGTCCACATAAATTTGAGAAAGCTCACTTACTTA	272	
DB	87	LysProPheAlaThrAspAsnGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu	106	
QY	273	AAGGCTATTCAAGAACAAATGATCGCTAAACGTCACAGTCCACAGTAAACGACGACTACTTTGAGGTC	332	
DB	107	LysAlaIleGlnLysGlnLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal	126	
QY	333	ATTGATTTTTCAGCGATGCCAACCTACTGATCGAATTTTTCGTAAGCGGACAT	452	
DB	127	IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp	146	
QY	393	AAAGATGTTGCGTTAACTTGGCCGACCAACTGTCCTCAAGAAATTTTTCGTAAGCGGACAT	452	
DB	147	LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis	166	
QY	453	GTGGCGGTTAGACCATATAAGAAAAACAATACAAAAACCAAGCGAAATCTGTTGATGTG	512	
DB	167	ValArgValArgProTyrLysGluLysProValGlnAsnGlnAlaLysSerValAspVal	186	
QY	513	GAATATACGTACAGTTTACTCCCTTAAACCTGATGACGATTTTCAGACCAGGTCTCAAA	572	
DB	187	LysTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys	206	
QY	573	GATCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATTACTA	632	
DB	207	AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu	226	
QY	633	GCTCAAGCACAAAGCATTTTAAACAAAAACCAACCCAGCGGTATACGATTTATGAACGTGAC	692	
DB	227	AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp	246	
QY	693	TCCTCAATCGTCACATGACAAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG	752	
DB	247	SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu	266	
QY	753	TTTACTTACCGTGTAAAAATCGGGAAACAAGCTTATAGGATCAATAAAAAATCTCGTCTG	812	
DB	267	PheThrTyrArgValLysAspArgGluGlnAlaTyrGlyIleAsnLysLysSerGlyLeu	286	
QY	813	AATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG	872	
DB	287	AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrIleLeuLysLysGly	306	
QY	873	GAAAGCGGTATGATCCCTTTGATCGCAGTCACATTGAACTGTTCCACCATCAATACGTT	932	
DB	307	GluSerProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal	326	
QY	933	GATGTCGATACCAACGAATTGCTAAAAAGTAGCAGCAGCTCTTAAACAGTAGTAGCGAAC	992	

Db 327 AspValAsnThrAsnGluLeuLeuLeuLeuSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
 QY TTAGACTTCAGAGATTATACGATCTCGTGTAAAGCGTAAACTACTCTACACAATCTC 1052
 D - - - - -
 Db 347 LeuAspPheArgAspLeuTyrAspProCysAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
 QY 1053 GATCGTTTGGTATTATGCGATATACCTTAACTGGAAAAAGTAGAGGATAATCACGATGAC 1112
 D - - - - -
 Db 367 AspAlaPheAspIleuMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspLys 386
 QY 1113 ACCAACCGTATCATACCGTTTATATGGCAGCGACCGAAGGAGAGAAATGCTACTAC 1172
 D - - - - -
 Db 387 AsnAsnArgIleValThrValTyrMetGlyLysArgProLysGlyAlaLysGlySerTyr 406
 QY 1173 CATTTAGCT 1181
 D - - - - -
 Db 407 HisLeuAla 409

RESULT 8

STRP_STRPY STANDARD; PRT; 440 AA.
 ID_P10520;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Streptokinase A precursor.
 GN Name=eka; OrderedLocusNames=SPY1979;
 OS Streptococcus pyogenes.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1314;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=SF130/13 / Serotype M1;
 RX MEDLINE=89160264; PubMed=2646590;
 RA Walter F., Siegel M., Malke H.;
 RT "Nucleotide sequence of the streptokinase gene from a Streptococcus
 pyogenes type 1 strain";
 RL Nucleic Acids Res. 17:1261-1261(1989).
 RN [2]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=SF370 / ATCC 700294 / Serotype M1;
 RX MEDLINE=21192684; PubMed=11296296; DOI=10.1073/pnas.071559398;
 RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
 RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
 RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
 RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.E.;
 RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
 CC -!- FUNCTION: This protein is not a protease, but it activates
 plasminogen by complexing with it. As a potential virulence
 factor, it is thought to prevent the formation of effective fibrin
 barriers around the site of infection, thereby contributing to the
 invasiveness of the cells.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use as long as its content is in no way modified and this statement is not
 removed.
 CC -----
 DR EMBL; X13399; CA311765.1; -; Genomic DNA.
 DR EMBL; AB006620; AAK34665.1; -; Genomic DNA.
 DR PIR; S02724; S02724.
 DR HSSP; Q53284; 1C4P.
 DR SMR; P10520; 63-398.
 DR InterPro; IPR004093; Staphylokinase.
 DR InterPro; IPR008124; Streptokinase.
 DR Pfam; PF02821; Staphylokinase; 3.
 DR PRINTS; PR01753; STREPKINASE.
 KW Complete proteome; Plasminogen activation; Signal; Virulence.
 FT SIGNAL 1 26
 FT CHAIN 27 440 Streptokinase A.

FT CONFLICT 163 163 L -> V (in Ref. 1).
 FT CONFLICT 345 345 R -> G (in Ref. 1).
 FT CONFLICT 373 373 D -> N (in Ref. 1).
 FT CONFLICT 428 428 D -> Y (in Ref. 1).
 FT CONFLICT 438 438 K -> N (in Ref. 1).
 SQ SEQUENCE 440 AA; 49924 MW; D6227EF040B758DB CRC64;
 Alignment Scores:
 Pred. No.: 8,76e-120 Length: 440
 Score: 1794.00 Matches: 341
 Percent Similarity: 93.99% Conservative: 19
 Best Local Similarity: 89.03% Mismatches: 23
 Query Match: 66.05% Indels: 0
 DB: 1 Gaps: 0
 US-09-940-235-9 (1-1541) x STRP_STRPY (1-440)
 QY 33 ATAGCTGGTCTCGAATGGCTACTAGATCGTCTCTCTTAATAACACGCAATGGTGT 92
 Db 27 IleAlaGlyTyrGlyTyrLeuProAspArgProIleAsnAsnSerGlnLeuValVal 46
 QY 93 AGCGTTGCTGTTACTGTTGAGGGGACGAATCAAGACATAGTCTTAAATTTTTTGAATC 152
 Db 47 SerMetAlaGlyIleValGluGlyThrAspLysLysValPheIleAsnPhePheGluIle 66
 QY 153 GATCTAACATCAGCACTGCTCATGAGGAGGAGACAGACGAGCTTAAGTCCAAATCA 212
 Db 67 AspLeuThrSerGlnProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 86
 QY 213 AAACCACTTGTCTACTGATGTCGCGCATGTCTACATAAACTTCAGAAAGCTGACTACTA 272
 Db 87 LysProPheAlaThrAspAsnGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu 106
 QY 273 AAGGCTATTCAAGAACAAATTGATCGTTAAGTCCACAGTAACGACGACTACTTTAGGTC 332
 Db 107 LysAlaIleGlnLysGlnLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal 126
 QY 333 ATTGATTTTGCAGCGATCGACCACTTACTGTCGAAACGGCAAGCTCTACTTTGCTGAC 332
 Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146
 QY 393 AAAGATGTTTCGGTAACTTCGCGACCACTGTCCAAAGATTTTTCCTTAACGGACAT 452
 Db 147 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuLysGlyHis 166
 QY 453 GTCGCGTTAGACCATATATAAGAAAAACCAATAACAAACCAACGAAATCTGTTGATGTG 512
 Db 167 ValArgValArgProTyrLysGluLysProValGlnAsnGlnAlaLysSerValAspVal 186
 QY 513 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGTCTCAAA 572
 Db 187 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 206
 QY 573 GATACCTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTA 632
 Db 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 226
 QY 633 GCTCAAGCAACAAGCAATTTTAAACAAAAACCAACCCAGGCTATACGATTTTATGAACGTGAC 692
 Db 227 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp 246
 QY 693 TCCTCAATGTCCTCATGACAAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 752
 Db 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
 QY 753 TTTACTTACCGTGTAAAATCGGGAACAAGCTTATAGGATCAATAAAAATCTGGTCTG 812
 Db 267 PheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsnProLysThrGlyLe 286
 QY 813 AATGAGAAATAACCAACACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAGGG 872
 Db 287 LysGluLysThrAsnAsnThrAspLeuValSerGluLysTyrTyrValLeuLysGlnGly 306

```
QY 873 GAAAGCCGTATGATCCCTTTGATCGAGTCACTTGAACCTGTTCAACATCAATAGCTT 932
Db 307 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326
QY 933 GATGTCGATACCAACGAATTTGCTAAAGAGGAGCAGCTCTTAACAGCTAGCGACGTAAAC 992
Db 327 AspValAsnThrAsnGlnLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
QY 993 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAAGCTACTCTCAACAATCTC 1052
Db 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
QY 1053 GATGCTTTTGGTATTATGACTATACCTTAACCTGGAAGAGTAGAGGATATACAGATGAC 1112
Db 367 AspAlaPheAspIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspLys 386
QY 1113 ACCAACCGTATCATACCTTTATATGGCAAGCGCCGAGAGAGATGCTAGCTAC 1172
Db 387 AsnAsnArgValValThrValTyrMetGlyLysArgProLysGlyAlaLysGlySerTyr 406
QY 1173 CATTTAGCT 1181
Db 407 HisLeuAla 409

RESULT 9
ID Q53ZX6_STRPY PRELIMINARY; PRT; 440 AA.
AC Q53ZX6.
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE Streptokinase.
GN Name=ska;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=86-779.
RX PubMed=14679231; DOI=10.1128/JB.186.1.110-121.2004;
RA Kalia A., Bessen D.E.;
RT "Natural selection and evolution of streptococcal virulence genes
involved in tissue-specific adaptations.";
RL J. Bacteriol. 186:110-121(2004);
DR EMBL, AY234128; AAP39948.1; -; Genomic_DNA.
KW Kinase.
SQ SEQUENCE 440 AA; 49924 MW; D6227EF040B758DB CRC64;

Alignment Scores:
Pred. No.: 8.76e-120 Length: 440
Score: 1794.00 Matches: 341
Percent Similarity: 93.9% Conservative: 19
Best Local Similarity: 89.03% Mismatches: 23
Query Match: 66.05% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-9 (1-1541) x Q53ZX6_STRPY (1-440)
QY 33 ATAGCTGGTCTGAATGGCTACTAGATCGTCTTCTGTAATAACAGCCAAATGGTTGTT 92
Db 27 IleAlaGlyTyrGlyTyrLeuProAspArgProProIleAsnAsnSerGlnLeuValVal 46
QY 93 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAATATTTTGAATC 152
Db 47 SerMetAlaGlyIleValGluGlyThrAspLysLysValPheIleAsnPheGluIle 66
QY 153 GATCTAACATCATGACCTCTCATGGAGGAAGACAGACAGCGCTTAAGTCCAAATCA 212
Db 67 AspLeuThrSerGlnProAlaHisGlyLysThrGluGlnGlyLeuSerProLysSer 86
QY 213 AAACCAATTTGCTACTGATAGTCGCGGATGTCATATAAATTCAGAAAGCTGACTTACTA 272
```

```
Db 87 LysProPheAlaThrAspAsnGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu 106
QY 273 AAGGCTATTCAAGAACAAATTCATCGCTAACCTGCACAGTAACAGACTACTTTGAGTGC 332
Db 107 LysAlaIleGlnLysGlnLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal 126
QY 333 ATTGATTTTGCACGATGCAACCATTTACTGATCGAAACGGCAGGCTTACTTTGCTGAC 392
Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146
QY 393 AAAGATGGTTGGTAACTTCCTGCCAGCCCACTGTCGAAGATTTTTCCTAAGCGGACAT 452
Db 147 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLysGlyHis 166
QY 453 GTGCGCGTTAGACCATATAAAGAAAAACAATACAAACCAACCAAGCGAATCTGTTGATCG 512
Db 167 ValArgValArgProTyrLysGluLysProValGlnAsnGlnAlaLysSerValAspVal 186
QY 513 GAATATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTCAGACCAAGTCTCAA 572
Db 187 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 206
QY 573 GATTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTACTA 632
Db 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 226
QY 633 GCTCAAGCAACAAGCATTTTAAACAAACCAACCCAGCGCTATACGATTTATGAACGTGAC 692
Db 227 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp 246
QY 693 TCCTCAATCGTCACTCATGACATGACATTTTCCGTGACGATTTTACCAATCGATCAAGAG 752
Db 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
QY 753 TTTACTTACCGTGTAAAAATTCGGGAACAAGCTTATAGATCAATAAAAAATCTGGTCTG 812
Db 267 PheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsnProLysThrGlyIle 286
QY 813 AATGAAGAAATAACAACACTGACTGATCTCTGAGAAATATTAGTCCTTAAAAAAGGG 872
Db 287 LysGluLysThrAsnAsnThrAspLeuValSerGluLysTyrTyrValLeuLysGlnGly 306
QY 873 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCAACATCAATAGCTT 932
Db 307 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326
QY 933 GATGTCGATACCAACGAATTTGCTAAAGAGGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 992
Db 327 AspValAsnThrAsnGlnLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
QY 993 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAAGCTACTCTCAACAATCTC 1052
Db 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
QY 1053 GATGCTTTTGGTATTATGACTATACCTTAACCTGGAAGAGTAGAGGATATACAGATGAC 1112
Db 367 AspAlaPheAspIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspLys 386
QY 1113 ACCAACCGTATCATACCTTTATATGGCAAGCGCCGAGAGAGATGCTAGCTAC 1172
Db 387 AsnAsnArgValValThrValTyrMetGlyLysArgProLysGlyAlaLysGlySerTyr 406
QY 1173 CATTTAGCT 1181
Db 407 HisLeuAla 409

RESULT 10
Q7X0Y2_STRPY
ID Q7X0Y2_STRPY PRELIMINARY; PRT; 440 AA.
AC Q7X0Y2.
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
```

DE Streptokinase.
GN Name=ska;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_taxid=1314;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=D306;
RX PubMed=14679231; DOI=10.1128/JB.186.1.110-121.2004;
RA Kalia A., Bessen D.E.;
RT "Natural selection and evolution of streptococcal virulence genes
involved in tissue-specific adaptations.";
RL J. Bacteriol. 186:110-121(2004).
DR EMBL; AV234136; AAP39956.1; -; Genomic_DNA.
DR HSP; Q53284; 1C4P.
DR SMR; Q7X0Y2; 63-398.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008243; F:plasminogen activator activity; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR InterPro; IPR004093; Staphylokinase.
DR InterPro; IPR008124; Streptokinase.
DR Pfam; PF02821; Staphylokinase; 3.
DR PRINTS; PR01753; STREPKINASE.
KW Kinase.
SQ SEQUENCE 440 AA; 50188 MW; 4CB42AC2A9062C2E CRC64;

Alignment Scores:

Pred. No.: 2,78e-119 Length: 440
Score: 1787.00 Matches: 342
Percent Similarity: 94.26% Conservative: 19
Best Local Similarity: 89.30% Mismatches: 22
Query Match: 65.80% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-9 (1-1541) x Q7X0Y2_STRPY (1-440)

QY 33 ATAGTGGTCTGTAAGTGGCTACTAGATCGTCTCTCTGTAATAACAGCAATGGTGT 92
Db 27 lleaAlaGlyTyrGlyTrpLeuProAspArgProIleAsnAsnSerGlnLeuVal 46
QY 93 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATAGTCTTAAATTTTGAATC 152
Db 47 SerMetAlaGlyIleValGluGlyThrAspLysLysValPheIleAsnPheGluIle 66
QY 153 GATCTAACATCAGACCTGCTCATGGAGGAAGACAGACGAGCTTAAGTCCAAATCA 212
Db 67 AspLeuThrSerGlnProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 86
QY 213 AAACCATTTGCTACTGATAGTGGCGGATGTCACATATAACTTGAGAAAGCTGACTTA 272
Db 87 LysProPheAlaThrAspAsnSerAlaMetProHisLysLeuGluLysAlaAspLeu 106
QY 273 AAGCGTATTCAAGAACAAATTGATCGCTAACGCTCCACAGTAACGACGACTACTTGA 332
Db 107 LysAlaIleGlnLysGlnLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal 126
QY 333 ATTGATTTTGAAGCGATGCAACCAATTACTGATCGAAACGGCAAGCTTACTTTGCTG 392
Db 127 lleaAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146
QY 393 AAAGATGTTGCTGTAACCTGCGGACCAACCTGCTCCAGAAATTTTGTACCGGCAT 452
Db 147 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuThrGlyHis 166
QY 453 GTGCCGCTTAGACCATATAAGAAACCAATACAAACCAACGAAATCTGTTGATGTG 512
Db 167 ValArgValArgProTyrLysGluLysProValGlnAsnGlnAlaLysSerValAspVal 186
QY 513 GAATATCTGTACGTTACTCCCTTAACCTTGATGACGATTTTCAGACAGGTCTCAAA 572
Db 187 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheLysProValLeuLys 206

QY 573 GATACTAAGCTATTGAAACACACTAGCTATCGGTGACACCATCATCTCAAGATTA 632
Db 207 AspThrLysLeuLeuLysLysLeuAlaIleGlyAspThrValThrSerGlnLeuLeu 226
QY 633 GCTCAAGCACAAAGCATTTTAAACAAACACCCAGGCTATACGATTTATGAACGTG 692
Db 227 AlaGlnAlaGlnSerIleLeuAsnGluSerHisProAspTyrThrIleTyrGluArgAsp 246
QY 693 TCCTCAATCGTCACTCATGACATATGATCTTCGTTACGATTTTACCAATGGATCAAG 752
Db 247 SerSerIleValThrHisAspLysAspIlePheArgThrIleLeuProMetAspGln 266
QY 753 TTTACTTACCGTGTATAAAATCGGAAACAGCTTTATAGGATCAATAAAAAATCTG 812
Db 267 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGly 286
QY 813 AATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGCTCTTAAAAA 872
Db 287 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLys 306
QY 873 GAAAGCCGTATGATCCCTTTGATCGCAGTCACCTGAACTGTTCAACCATCAATAG 932
Db 307 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysVal 326
QY 933 GATGTCGATACCAACGAATTTCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCAAC 992
Db 327 AspValAsnThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
QY 993 TTAGACTTCAGAGATTTATACGATCTCTGTAAGCTAACTACTCTTACAAACATCT 1052
Db 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsn 366
QY 1053 GATGCTTTGGTATTATGACTATACCTTAACCTGGAAGTAGAGGATATACGATG 1112
Db 367 AspAlaPheAspIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAsp 386
QY 1113 ACCAACGCTATACACCGTTTATATGGCAACGACCGCAAGAGAGAGATGCTAGTAC 1172
Db 387 AsnAsnArgValValThrValTyrMetGlyLysArgProLysGlyAlaLysGlySer 406
QY 1173 CATTAGCT 1181
Db 407 HisLeuAla 409
RESULT 11
Q7X0Y7_STRPY
ID Q7X0Y7_STRPY PRELIMINARY; PRT; 440 AA.
AC Q7X0Y7;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Streptokinase.
GN Name=ska;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_taxid=1314;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=IRP112;
RX PubMed=14679231; DOI=10.1128/JB.186.1.110-121.2004;
RA Kalia A., Bessen D.E.;
RT "Natural selection and evolution of streptococcal virulence genes
involved in tissue-specific adaptations.";
RL J. Bacteriol. 186:110-121(2004).
DR EMBL; AV234130; AAP39950.1; -; Genomic_DNA.
DR HSP; Q53284; 1C4P.
DR SMR; Q7X0Y7; 63-398.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008243; F:plasminogen activator activity; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR InterPro; IPR004093; Staphylokinase.
DR InterPro; IPR008124; Streptokinase.

DR Pfam: PF02821; Staphylokinase; 3.
DR PRINTS: PR01753; STREPKINASE.
KW Kinase.
SQ SEQUENCE 440 AA; 50160 MW; 1C5AF07907BC7AC8 CRC64;

Alignment Scores:
Pred. No.: 3.3e-117 Length: 440
Score: 1758.00 Matches: 338
Percent Similarity: 93.21% Conservative: 19
Best Local Similarity: 88.25% Mismatches: 26
Query Match: 64.73% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-9 (1-1541) x Q7X0Y7_STRPY (1-440)

QY 33 ATAGCTGGTCCCTGAATGGCTACTAGATCGCTCTTCTGTAATAACAGCCAAATGGTTGTT 92
DB 27 IleAlaGlyTyrGlyTyrLeuLeuAspArgProValAsnAsnSerGlnLeuVal 46

QY 93 AGCGTTGCTGACTCTGTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 152
DB 47 SerMetAlaGlyLeuValGluGlyThrAspLysLysValPheIleAsnPheGluLe 66

QY 153 GATCTAATCATGACCTGCTCATGAGGAGAAACAGACAGCAAGCTTAATCCAAATCA 212
DB 67 AspLeuThrSerGlnProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 86

QY 213 AAACCATTTGCTACTGATGCGCGGATGTCATATAACTTCAGAAAGCTGACTTACTA 272
DB 87 LysProPheAlaThrAsnSerSerAlaMetProHisLysLeuGluLysAlaAspLeuLeu 106

QY 273 AAGGCTATTCAAGAACAAATGATCGCTTAACGTCACAGTAACGACGACTACTTTGAGTC 332
DB 107 LysAlaIleGlnGluArgLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal 126

QY 333 ATTGATTGTCGAAGCGATGCAACCATTAATGATCGAAACCGCAAGGTCTACTTTGCTGAC 392
DB 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146

QY 393 AAAGATGGTTCGGTAACCTTCGCGACCACTGTCACAGAAATTTTGTAAACGGACAT 452
DB 147 LysAspAspSerValThrLeuProThrGlnProValGlnGluPheLeuLeuArgGlyHis 166

QY 453 GTGCGGTGAGACCATATAAGAAAAACCAATACAAACCAACGAAATCTGTTGATGTG 512
DB 167 ValArgValArgProTyrLysGluLysProIleGlnThrProAlaLysSerValAspVal 186

QY 513 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTACAGCAGGTCTCAAA 572
DB 187 ArgTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProValLeuLys 206

QY 573 GATCTAGCTATTGAAACACTAGTATCGGTGACACCATCATCATCTCAAGAAATTACTA 632
DB 207 AsnThrLysLeuLysThrLeuAlaIleGlyGlyThrValThrSerGlnGluLeuLeu 226

QY 633 GCTCAAGCAACAGCATTTTAAACAAACCAACCCAGGCTATACGATTTATGACCTGAC 692
DB 227 AlaGlnAlaGlnSerIleLeuAsnGluSerHisProAspTyrThrIleTyrGluArgAsp 246

QY 693 TCCTCAATCGTCACTCATGACATGATGATTTTCCTGATGATTTTACCAATGGATCAAGAG 752
DB 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266

QY 753 TTACTTACCGGTGTTAAATCGGAAACAGCTTATAGGATCAATAAAAAATCTGCTGTC 812
DB 267 PheThrTyrHisIleLysAspArgGluGlnAlaTyrGlyIleAsnLysLysSerGlyGln 286

QY 813 AATGAAGAAATAAACHACACTGACCTGATCTCTGAGAAATATTAAGTCTTAAAAAGGG 872
DB 287 GluGluLysThrAsnAsnThrAspLeuIleSerGluLysTyrValLeuLysLysGly 306

QY 873 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTTCCACCATCAATAGTT 932
DB 873 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTTCCACCATCAATAGTT 932

DB 307 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleAsnTyrVal 326

QY 933 GATGTCGATACCAACGAATTCCTAAAGAGTCCAGAGCTCTTTAAACAGCTAGCGAACTAAC 992
DB 327 AspValAsnThrAsnLysLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346

QY 993 TTAGACTTCAGAGATTTATACGATCCTCGTGATAGGCTAAAGCTAACTCTACAACTCTC 1052
DB 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366

QY 1053 GATGCTTTTGGTATTATGGAATATACCTTAACCTGGAAGAGTAGAGGATAATCAGATGAC 1112
DB 367 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspLys 386

QY 1113 ACCAACCGTATCAACCGTTTATATGGCAAGCGACCCGAGAGAGAGAGAACTGCTAGCTAC 1172
DB 387 AsnAsnArgValValThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 406

QY 1173 CATTTAGCT 1181
DB 407 HisLeuAla 409

RESULT 12
Q5X9T6_STRP6
ID Q5X9T6_STRP6 PRELIMINARY; PRT; 440 AA.
AC Q5X9T6;
DC 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Streptokinase (EC 3.4.11.1).
GN OrderedLocustNames=M6_Spy1692;
OS Streptococcus pyogenes (serotype M6).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=301450;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MGAS10394;
RX PubMed=15272401; DOI=10.1086/422697;
RA Banks D.J., Porcella S.F., Barbican K.D., Beres S.B., Phillips L.E.,
RA Voyich J.M., DeLeo F.R., Martin J.M., Somerville G.A., Musser J.M.;
RT "Progress toward characterization of the group A Streptococcus
metagenome: complete genome sequence of a macrolide-resistant serotype
M6 strain."
RL J. Infect. Dis. 190:727-738 (2004).
DR EMBL; CP000003; AAT87827.1; -; Genomic_DNA.
DR SMR; Q5X9T6; 63-398.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008243; F:plasmaenogen activator activity; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR InterPro; IPR004093; Staphylokinase.
DR InterPro; IPR008124; Streptokinase.
DR Pfam; PF02821; Staphylokinase; 3.
DR PRINTS; PR01753; STREPKINASE.
KW Complete proteome; Hydrolase; Kinase.
SQ SEQUENCE 440 AA; 50160 MW; 1C5AF07907BC7AC8 CRC64;

Alignment Scores:
Pred. No.: 3.3e-117 Length: 440
Score: 1758.00 Matches: 338
Percent Similarity: 93.21% Conservative: 19
Best Local Similarity: 88.25% Mismatches: 26
Query Match: 64.73% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-9 (1-1541) x Q5X9T6_STRP6 (1-440)

QY 33 ATAGCTGGTCCCTGAATGGCTACTAGATCGCTCTTCTGTAATAACAGCCAAATGGTTGTT 92
DB 27 IleAlaGlyTyrGlyTyrLeuLeuAspArgProValAsnAsnSerGlnLeuVal 46

QY 93 AGCGTTGCTGACTCTGTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 152
DB 47 SerMetAlaGlyLeuValGluGlyThrAspLysLysValPheIleAsnPheGluLe 66

QY 153 GATCTAATCATGACCTGCTCATGAGGAGAAACAGACAGCAAGCTTAATCCAAATCA 212
DB 67 AspLeuThrSerGlnProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 86

QY 213 AAACCATTTGCTACTGATGCGCGGATGTCATATAACTTCAGAAAGCTGACTTACTA 272
DB 87 LysProPheAlaThrAsnSerSerAlaMetProHisLysLeuGluLysAlaAspLeuLeu 106

QY 273 AAGGCTATTCAAGAACAAATGATCGCTTAACGTCACAGTAACGACGACTACTTTGAGTC 332
DB 107 LysAlaIleGlnGluArgLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal 126

QY 333 ATTGATTGTCGAAGCGATGCAACCATTAATGATCGAAACCGCAAGGTCTACTTTGCTGAC 392
DB 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146

QY 393 AAAGATGGTTCGGTAACCTTCGCGACCACTGTCACAGAAATTTTGTAAACGGACAT 452
DB 147 LysAspAspSerValThrLeuProThrGlnProValGlnGluPheLeuLeuArgGlyHis 166

QY 453 GTGCGGTGAGACCATATAAGAAAAACCAATACAAACCAACGAAATCTGTTGATGTG 512
DB 167 ValArgValArgProTyrLysGluLysProIleGlnThrProAlaLysSerValAspVal 186

QY 513 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTACAGCAGGTCTCAAA 572
DB 187 ArgTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProValLeuLys 206

QY 573 GATCTAGCTATTGAAACACTAGTATCGGTGACACCATCATCATCTCAAGAAATTACTA 632
DB 207 AsnThrLysLeuLysThrLeuAlaIleGlyGlyThrValThrSerGlnGluLeuLeu 226

QY 633 GCTCAAGCAACAGCATTTTAAACAAACCAACCCAGGCTATACGATTTATGACCTGAC 692
DB 227 AlaGlnAlaGlnSerIleLeuAsnGluSerHisProAspTyrThrIleTyrGluArgAsp 246

QY 693 TCCTCAATCGTCACTCATGACATGATGATTTTCCTGATGATTTTACCAATGGATCAAGAG 752
DB 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266

QY 753 TTACTTACCGGTGTTAAATCGGAAACAGCTTATAGGATCAATAAAAAATCTGCTGTC 812
DB 267 PheThrTyrHisIleLysAspArgGluGlnAlaTyrGlyIleAsnLysLysSerGlyGln 286

QY 813 AATGAAGAAATAAACHACACTGACCTGATCTCTGAGAAATATTAAGTCTTAAAAAGGG 872
DB 287 GluGluLysThrAsnAsnThrAspLeuIleSerGluLysTyrValLeuLysLysGly 306

QY 873 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTTCCACCATCAATAGTT 932
DB 873 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTTCCACCATCAATAGTT 932

QY	33	ATAGCTGTCCTGANTGCTACTAGATCGTCTCTCTGTAATAACAGCCAAATGGTGT	92
Db	27	IleAlaGlyTyrGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal	46
QY	93	AGCGTTGCTGGTACTGTTGAGGGGACGATCAAGACATAGTCTATAATTTTTTGAATC	152
Db	47	SerMetAlaGlyIleValGluGlyThrAspLysLysValPheIleAsnPheGluIle	66
QY	153	GATCTAACATCACGACCTGCTCATCGAGGAAAGACAGACGCAAGGCTTAAGTCCAAATCA	212
Db	67	AspLeuThrSerGlnProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer	86
QY	213	AAACCATTTGCTACTGATAGTGGCGCGATGTCACATAAACTTGAGAACTGACTTACTA	272
Db	87	LysProPheAlaThrAspAsnGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu	106
QY	273	AAGCGTATTCAAGACAATCGATCGCTACGTCACAGTCCACAGTAAACGACGACTCTT	332
Db	107	LysAlaIleGlnGluArgLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal	126
QY	333	ATTGATTTTGAAGCGATGCAACCACTATCTGATCGAAACGGAGTCTACTTTGCTGAC	392
Db	127	IleAspPheAlaSerAspAlaThrIleThrAspArgAspAsnIleTyrPheAlaAsn	146
QY	393	AAAGATGTTCCGTAACCTTGCCGACCCAACTGTGCCAAGAAATTTTGTAAAGCGGACAT	452
Db	147	GlnAspGlySerValThrLeuProThrGlnProIleGlnGlnPheLeuLeuArgGlyHis	166
QY	453	GTGCGCGTTAGACCATATAAGAAAAACCAATACAAAACCAACGCGAACTCTGTGATG	512
Db	167	ValArgValArgProTyrLysGluLysProIleGlnThrProAlaLysSerValAspIle	186
QY	513	GAATATCTGTACAGTTTACTTCCCTTAAACCTGTATGACGATTTTCAGACGAGTCTCAAA	572
Db	187	ArgTyrThrValGlnPheThrProLeuAsnProAspAspAspPheLysProValLeuLys	206
QY	573	GATACTAAGCTATTGAAACACTAGCTATCGGTGACCAACCATCACATCTCTCAAGAATTACTA	632
Db	207	AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu	226
QY	633	GCTCAGCACAAGCATTTTAAACAAAAACCAACCCAGCGGTATACGATTTATGAACGTGAC	692
Db	227	AlaGlnAlaGlnSerIleLeuAsnGluSerHisProAspTyrThrIleTyrGluArgAsp	246
QY	693	TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGATCAAGAG	752

Db 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProThrAspGlnGlu 266
|||||
Qy 753 TTTACTTACCGTGTAAATAATCGGAACAAAGCTTATAGGATCAATAAATAATCTGCTCTG 812
|||||
Db 267 PheThrTyHisValLysAsnArgGluGlnAlaTyGlnAsnAspAsnLysThrGlyLeu 286
|||||
Qy 813 AATGAAGAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 872
|||||
Db 287 LysLysGluThrLysAsnThrAspLeuIleSerGluLysTyTyrlleLeuLysGly 306
|||||
Qy 873 GAAAAGCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCAACATCAAATACGTT 932
|||||
Db 307 GluLysProTyAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyVal 326
|||||
Qy 933 GATGTCGATACCAACGAATTGCTTAAAGTGGACGCTCTTAACAGCTAGCGACCTAAC 992
|||||
Db 327 AspValAspThrLysAspLeuLysSerGluGlnLeuThrAlaSerGluArgAsn 346
|||||
Qy 993 TTAGACTTCAGAGATTATACGATCCTCGTGAAGGCTAAACTACTCTACAAATCTC 1052
|||||
Db 347 LeuAspPheArgAspLeuTyAspProArgAspLysAlaLysLeuLeuTyrlanAsnLeu 366
|||||
Qy 1053 GATGCTTTGGTATTATGACTATATACCTTAAGTGGAAAAAGTAGAGGATAATCACGATGAC 1112
|||||
Db 367 AspAlaPheAspIleMetAspTyThrLeuThrGlyLysValGluAspAsnHisAspLys 386
|||||
Qy 1113 ACCAACCGTATCATACCTTTATGCGCAAGCGACGAGGAGAGATGCTACTAC 1172
|||||
Db 387 AsnAsnArgValThrValTyMetGlyLysArgProLysGlyAlaLysGlySerTy 406
|||||
Qy 1173 CATTTAGCT 1181
|||||
Db 407 HisLeuAla 409

RESULT 15

Q7X0Y0 STRPY PRELIMINARY; PRT; 440 AA.
ID Q7X0Y0- STRPY PRELIMINARY; PRT; 440 AA.
AC Q7X0Y0-
DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Streptokinase.
GN Name=ska;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=d633;
RX PubMed=14679231; DOI=10.1128/JB.186.1.110-121.2004;
RA Kalia A., Bessen D.E.;
RT "Natural selection and evolution of streptococcal virulence genes involved in tissue-specific adaptations.";
RL J. Bacteriol. 186:110-121(2004).
DR EMBL; AY234138; AAP39958.1; -; Genomic_DNA.
DR HSSP; Q53284; 1C4P.
DR SMR; Q7X0Y0; 63-398.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008243; F:plasmaenogen activator activity; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR InterPro; IPR004093; Staphylokinase.
DR InterPro; IPR008124; Streptokinase.
DR Pfam; PF02821; Staphylokinase; 3.
DR PRINTS; PR01753; STREPKINASE.
KW Kinase.
SQ SEQUENCE 440 AA; 50143 MW; 4ACB2C29349D680C CRC64;

Alignment Scores:

Pred. No.: 3,95e-114 Length: 440
Score: 1715.00 Matches: 325
Percent Similarity: 91.91% Conservative: 27

Best Local Similarity: 84.86% Mismatches: 31
Query Match: 63.14% Indels: 0
DB: 2 Gaps: 0
US-09-940-235-9 (1-1541) x Q7X0Y0_STRPY (1-440)
Qy 33 ATAGCTGGTCTGAATGGCTACTAGATCGCTCTTGTAAATAACGCCAATGGTTGTT 92
|||||
Db 27 IleAlaGlyProGluTrpLeuLeuGlyArgProProValAsnAsnSerGlnLeuVal 46
|||||
Qy 93 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAAGCTTTAAATTTTGAATC 152
|||||
Db 47 SerMetAlaGlyIleValGluGlyThrAspLysValPheIleAsnPheGluIle 66
|||||
Qy 153 GATCTAACATCACGACCTGCTCATGAGGAGAAACAGACAGCAAGCTTAAGTCCAAATCA 212
|||||
Db 67 AspLeuThrSerGlnProAlaHisGlyGlyThrGlnGlnLysLeuSerProLysSer 86
|||||
Qy 213 AAACCATTTTCTACTGATGCGCGGATGTCATAAACTTGAGAAAGCTGACTTACTA 272
|||||
Db 87 LysProPheAlaThrAspAsnGlyAlaMetProHisLysLeuGluLysAlaAspLeu 106
|||||
Qy 273 AAGGCTATTCAAGAACAAATGATCGCTTAACGTCACAGTAACGACGACTACTTTGAGCT 332
|||||
Db 107 LysAlaIleGlnGluArgLeuIleAlaAsnValHisSerAsnAspGlyTyThrPheGluVal 126
|||||
Qy 333 ATTGATTTTCAAGCGATGCAACATTTACTGTCGAAACCGCAAGGCTCTACTTTGCTGAC 392
|||||
Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAspAsnIleTyThrPheAlaAsn 146
|||||
Qy 393 AAAGATGGTTCGGTAACTTTGCCGACCCCAACCTGTCCAAGAAATTTTGTAAAGCGGACAT 452
|||||
Db 147 GlnAspGlySerValThrLeuProThrGlnProIleGlnPheLeuLeuArgGlyHis 166
|||||
Qy 453 GTCCGCTTAGACCATATAAAGAAAAACCAATACAAACCAACGCAAACTGTGTGATGTG 512
|||||
Db 167 ValArgValArgProTyLysGluLysProIleGlnThrProAlaLysSerValAspIle 186
|||||
Qy 513 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAA 572
|||||
Db 187 ArgTyThrValGlnPheThrProLeuAsnProAspAspPheLysProValLeuLys 206
|||||
Qy 573 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTA 632
|||||
Db 207 AspThrLysLeuLysThrLeuAlaIleGlyAsnThrIleThrSerGlnGlnLeuLeu 226
|||||
Qy 633 GCTCAAGCACAAAGCATTTTAAACAAAAACCCAGGCTATACGATTTATGAACGTCGAC 692
|||||
Db 227 AlaGlnAlaGlnSerIleLeuAsnGlnSerHisProAspTyThrIleTyGluArgAsp 246
|||||
Qy 693 TCCTCAATCTGCTCATGCAATGACATTTTCGGTACGATTTTTCACATGGATCAAGAG 752
|||||
Db 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProThrAspGlnGlu 266
|||||
Qy 753 TTTTACTTACCGTGTAAAAATCGGAACAGCTTATAGGATCAATAAATAATCTGGCTG 812
|||||
Db 267 PheThrTyHisValLysAsnArgGluGlnAlaTyrlLysAlaAsnSerLysThrAspIle 286
|||||
Qy 813 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 872
|||||
Db 287 LysGluLysThrAsnAsnThrAspLeuIleSerGluLysTyTyValLeuLysLysGly 306
|||||
Qy 873 GAAAAGCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCAACATCAAATACGTT 932
|||||
Db 307 GluGluProTyAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyVal 326
|||||
Qy 933 GATGTCGATACCAACGAATTGCTTAAAGTGGACGCTCTTAACAGCTAGCGACCTAAC 992
|||||
Db 327 AspValAspThrLysAlaLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
|||||
Qy 993 TTAGACTTCAGAGATTATACGATCCTCGTGAAGGCTAAACTACTCTACAAATCTC 1052
|||||
Db 347 LeuAspPheArgAspLeuTyAspProArgAspLysAlaLysLeuLeuTyrlanAsnLeu 366


```

QY 1053 GATGCTTTGGTATTATGGACTATACCTTAACCGAAAAGTAGAGGATAATCAGCATCAC 1112
Db |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
367 AspAlaPheAspIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspLys 386
QY 1113 ACCAAGCGTATCATACCGTTTATATGGGCAAGCGACCCGAAAGAGAGAAATGCTAGCTAC 1172
Db |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
387 AsnAsnArgValThrValTyrMetGlyLysArgProLysGlyAlaLysGlySerTyr 406
QY 1173 CATTAGCT 1181
Db |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
407 HisLeuAla 409

```

Search completed: January 28, 2006, 02:33:35
Job time : 183.071 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2006, 10:53:18 ; Search time 8444.12 Seconds
(without alignments)
11181.386 Million cell updates/sec

Title: US-09-940-235-10
Perfect score: 1661
Sequence: 1 gaacacccgcgcctagcc.....gaataagctgtaccatctaa 1661

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues
Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

1: gb_ba.*
2: gb_in.*
3: gb_env.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pr.*
9: gb_ro.*
10: gb_ste.*
11: gb_sy.*
12: gb_un.*
13: gb_vi.*
14: gb_htg.*
15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1150	69.2	2385	6	AR143998 Sequence
2	1148.6	69.2	1242	6	AR144000 Sequence
3	1148.6	69.2	1401	6	I05204 Sequence 5
4	1148.6	69.2	2568	1	STRSKC
5	1148.6	69.2	2568	6	A04926 Streptococ
6	1148.6	69.2	2568	6	A04926 S.equisimil
7	1143.8	68.9	2568	1	SEDEXB
8	1136.2	68.4	7057	6	E00522 DNA fragmen
9	1134.6	68.3	2566	6	CQ97820 Sequence
10	1127.8	67.9	1245	1	S46536
11	1127.6	67.9	1257	6	A20015 SEQ ID NO:
12	1127.6	67.9	1257	6	I13203 Sequence 25
13	1126.2	67.8	1407	6	E01413 DNA sequenc
14	1126.2	67.8	1512	6	A20016 SEQ ID NO:
15	1126.2	67.8	1512	6	I13204 Sequence 27
16	1124.6	67.7	1317	6	A20009 SEQ ID NO:
17	1124.6	67.7	1317	6	I13197 Sequence 18
18	1124.6	67.7	1335	6	A20006 SEQ ID NO:

19 1124.6 67.7 1335 6 I13194
20 1124.6 67.7 1458 6 A20027
21 1124.6 67.7 1458 6 I13215
22 1124.6 67.7 1467 6 A20030
23 1124.6 67.7 1467 6 I13218
24 1124.6 67.7 2588 6 A20021
25 1124.6 67.7 2589 6 I13209
26 1116.6 67.2 1473 1 SGSKG
27 1115 67.1 1311 1 A2368335
28 1096.6 66.0 1122 6 ARI75893
29 1096.6 66.0 1122 6 AX030317
30 1096.6 66.0 1158 6 ARI75894
31 1096.6 66.0 1158 6 AX030318
32 1089.2 65.6 1209 6 ARI75891
33 1089.2 65.6 1209 6 AX030315
34 1089.2 65.6 1245 6 ARI75892
35 1089.2 65.6 1245 6 AX030316
36 1080.4 65.0 2253 6 I13213
37 1078.8 64.9 1119 6 I13206
38 1078.6 64.9 2252 6 A20025
39 1066.8 64.2 1118 6 A20018
40 1012.6 61.0 1323 1 A234129
41 1012.6 61.0 50354 1 AE014169
42 1012.6 61.0 110000 1 BA000034_16
43 1012.6 61.0 110000 1 BA000034_17
44 1011 60.9 1323 1 A234137
45 1004.6 60.5 1323 1 A234136

ALIGNMENTS

RESULT 1
LOCUS ARI43998 2385 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 1 from patent US 6210667.
ACCESSION ARI43998
VERSION ARI43998.1 GI:15105865
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2385)
AUTHORS Reed,G.L.
TITLE Bacterial fibrin-dependent plasminogen activator
JOURNAL Patent: US 6210667-A 1 03-APR-2001;
FEATURES Location/Qualifiers
source 1. 2385
/organism="unknown"
/mol_type="unassigned DNA"

Query Match 69.2%; Score 1150; DB 6; Length 2385;
Best Local Similarity 99.6%; Pred. No. 8.3e-292;
Matches 1153; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 180 ATGATTCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCAATTGGTT 239
Db 1141 AGGATTCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCAATTAGTT 1200
Qy 240 GTTAGCGTCTGCTGCTACTGTTGAGGGGACGATCAAGACATTAAGTCTTAATTTTGA 299
Db 1201 GTTAGCGTCTGCTGCTACTGTTGAGGGGACGATCAAGACATTAAGTCTTAATTTTGA 1260
Qy 300 ATCGATCTAACATCAACACCTGCTCATGAGGAGGAGACAGCAAGCTTAAGTCCAAAA 359
Db 1261 ATCGATCTAACATCAACACCTGCTCATGAGGAGGAGACAGCAAGCTTAAGTCCAAAA 1320
Qy 360 TCAAAACCAATTTGCTACTGATAGTGGCGCGATGTACATAACTTTGAGAAGCTGACTTA 419
Db 1321 TCAAAACCAATTTGCTACTGATAGTGGCGCGATGTACATAACTTTGAGAAGCTGACTTA 1380
Qy 420 CTAAAGGCTATTCAAGAACAAATTGATCGCTAAACGTCACAGTACACGACTACTTTTGAG 479

1381 CTAAAGGCTATTCAAGAACTTTGATCGTTAACTCCACAGTAAACGACGACTACTTTGAG 1440
480 GTCAATGATTTGCAAGCGATGCAACCATTAATGATCGAAACGGAAGGCTCTACTTTGCT 539
1441 GTCAATGATTTGCAAGCGATGCAACCATTAATGATCGAAACGGAAGGCTCTACTTTGCT 1500
540 GACAAGATGGTTCCGTAACCTTTGCCGACCCCAACCTGTCACAAATTTTTCGCTAAAGCGGA 599
1501 GACAAGATGGTTCCGTAACCTTTGCCGACCCCAACCTGTCACAAATTTTTCGCTAAAGCGGA 1560
600 CATGTGCGGTTAGACCATATAAAGAAACCAATACAAAACCAAGCGAAATCTGTTGAT 659
1561 CATGTGCGGTTAGACCATATAAAGAAACCAATACAAAACCAAGCGAAATCTGTTGAT 1620
660 GTGGAATATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTC 719
1621 GTGGAATATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTC 1680
720 AAGATATCTAGCTATTGAACACATAGCTATCGGTGACACCATCATCTCAAGAAATTA 779
1681 AAGATATCTAGCTATTGAACACATAGCTATCGGTGACACCATCATCTCAAGAAATTA 1740
780 CTAGCTCAAGCAAAAGCAATTTTAAACAAAACCAACCCAGGCTATACGATTTTATGAACGT 839
1741 CTAGCTCAAGCAAAAGCAATTTTAAACAAAACCAACCCAGGCTATACGATTTTATGAACGT 1800
840 GACTCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGGAATCAA 899
1801 GACTCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGGAATCAA 1860
900 GAGTTTACTTACCGTGTAAATAATCGGGAACAAGCTTATAGGATCAATAAAAAATCTGGT 959
1861 GAGTTTACTTACCGTGTAAATAATCGGGAACAAGCTTATAGGATCAATAAAAAATCTGGT 1920
960 CTGAATGAAGAAATAAACAACCTGATCTCTGAGAAATATTACGTTCCTTAAAAAA 1019
1921 CTGAATGAAGAAATAAACAACCTGATCTCTGAGAAATATTACGTTCCTTAAAAAA 1980
1020 GGGGAAAAGCGGTATGATCCCTTTGATCGAGTCACTTGAAACTGTTTCCACATCAATATC 1079
1981 GGGGAAAAGCGGTATGATCCCTTTGATCGAGTCACTTGAAACTGTTTCCACATCAATATC 2040
1080 GTTGATGTCGATACCAACGAATTTCTAAAGTGAAGAGCTCTTAAACAGCTAGCGAACGT 1139
2041 GTTGATGTCGATACCAACGAATTTCTAAAGTGAAGAGCTCTTAAACAGCTAGCGAACGT 2100
1140 AACTTTAGACTTCAGAGATTTATACGATCTCTGATGATGAGCTAACTCTTCAACAAT 1199
2101 AACTTTAGACTTCAGAGATTTATACGATCTCTGATGATGAGCTAACTCTTCAACAAT 2160
1200 CTGATGCTTTTGGTATTATGGAATATACCTTAACTGGAAGTATAGGATATACAGAT 1259
2161 CTGATGCTTTTGGTATTATGGAATATACCTTAACTGGAAGTATAGGATATACAGAT 2220
1260 GACACCAACCGTATCATACCGTTTATGCGCAAGCGACCCGAGAGAGAAATGCTAGC 1319
2221 GACACCAACCGTATCATACCGTTTATGCGCAAGCGACCCGAGAGAGAAATGCTAGC 2280
1320 TATCATTTAGCCGGTGGT 1337
2281 TATCATTTAGCCCTATGAT 2298

RESULT 2
AR144000
LOCUS AR144000
DEFINITION Sequence 5 from patent US 6210667.
ACCESSION AR144000
VERSION AR144000.1 GI:15105867
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

Unclassified.
REFERENCE 1 (bases 1 to 1242)
AUTHORS Reed,G.L.
TITLE Bacterial fibrin-dependent plasminogen activator
JOURNAL Patent: US 6210667-A 5 03-APR-2001;
FEATURES Location/Qualifiers
source 1. .1242
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 69.2%; Score 1148.6; DB 6; Length 1242;
Best Local Similarity 99.7%; Pred. No. 1.9e-291;
Matches 1151; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 183 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCAATTTGGTTGTT 242
DB 1 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCAATTTAGTTGTT 60
QY 243 AGCGTTCTGCTACTGTTGAGGGGACGAATCAAGACATTAAGCTTAAATTTTGAATC 302
DB 61 AGCGTTCTGCTACTGTTGAGGGGACGAATCAAGACATTAAGCTTAAATTTTGAATC 120
QY 303 GATCTAACATCAACGACCTGCTCATGGAGGAAAGACAGCAAGGCTTAAAGTCCAAATCA 362
DB 121 GATCTAACATCAACGACCTGCTCATGGAGGAAAGACAGCAAGGCTTAAAGTCCAAATCA 180
QY 363 AAACCAATTTGCTACTGATAGTGGCGGATGTCAATAACTTGAGAAAGCTGACTTACTA 422
DB 181 AAACCAATTTGCTACTGATAGTGGCGGATGTCAATAACTTGAGAAAGCTGACTTACTA 240
QY 423 AAGGCTATTCAAGAACTGATGCTTAACTGATCGTCAACAGTAAACGAGCTACTTTGAGTGC 482
DB 241 AAGGCTATTCAAGAACTGATGCTTAACTGATCGTCAACAGTAAACGAGCTACTTTGAGTGC 300
QY 483 ATTGATTTTCCAGCGATGCAACCATTAATGATCGTCAACAGTAAACGAGCTACTTTGAGTGC 542
DB 301 ATTGATTTTCCAGCGATGCAACCATTAATGATCGTCAACAGTAAACGAGCTACTTTGAGTGC 360
QY 543 AAAGATGGTTCCGTTAACTTTGCCGACCCCAACCTGTCCAAGAAATTTTGTAAAGCGGACAT 602
DB 361 AAAGATGGTTCCGTTAACTTTGCCGACCCCAACCTGTCCAAGAAATTTTGTAAAGCGGACAT 420
QY 603 GTGCGGCTTACACATATAAAGAAACCAATACAAAACCAAGCGAAATCTGTTGATGTC 662
DB 421 GTGCGGCTTACACATATAAAGAAACCAATACAAAACCAAGCGAAATCTGTTGATGTC 480
QY 663 GAATATATCTGTACAGTTTACTTCCCTTAAACCCCTGATGACGATTTTCCAGACAGGCTCTCAA 722
DB 481 GAATATATCTGTACAGTTTACTTCCCTTAAACCCCTGATGACGATTTTCCAGACAGGCTCTCAA 540
QY 723 GATATCTAAGCTATTGAAACACACTAGCTATCGGTGACACCATCATCTCAAGAAATTA 782
DB 541 GATATCTAAGCTATTGAAACACACTAGCTATCGGTGACACCATCATCTCAAGAAATTA 600
QY 783 GCTCAAGCAACAAGCATTTTAAACAAAACCAACCCAGGCTATACGATTTATGAACGTCAC 842
DB 601 GCTCAAGCAACAAGCATTTTAAACAAAACCAACCCAGGCTATACGATTTATGAACGTCAC 660
QY 843 TCCTCAATCTGCTCATGATGACATTTTCCGTCAGATTTTCCCAATGATCAAGCAAG 902
DB 661 TCCTCAATCTGCTCATGATGACATTTTCCGTCAGATTTTCCCAATGATCAAGCAAG 720
QY 903 TTTTACTTACCGTGTAAAAATCGGGAACAAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 962
DB 721 TTTTACTTACCGTGTAAAAATCGGGAACAAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 780
QY 963 AATGAAGAAATAAACAACCTGATCTCTGAGAAATATTACGCTCTTAAAGAGG 1022
DB 781 AATGAAGAAATAAACAACCTGATCTCTGAGAAATATTACGCTCTTAAAGAGG 840
QY 1023 GAAAAGCGGTATGATCCCTTTGATCGGAGTCACTTGAACCTGTTCAACATCAATACGTT 1082

```
Db      841  GAAAGCGGTATGATCCCTTTGATCGCAGTCATCTTGAAACTGTTTCAACATCAAAATACGTT 900
Qy      1083  GATGTCGATACCAACGAATTGCTAAAGATGAGCAGCTCTTTAAACAGCTAGCGAAACGTAAC 1142
Db      901  GATGTCGATACCAACGAATTGCTAAAGATGAGCAGCTCTTTAAACAGCTAGCGAAACGTAAC 960
Qy      1143  TTAGCTTCAGAGATTATACGATCTCTGCTGATGAAGCTTAACTTCTTACAAACATCTC 1202
Db      961  TTAGCTTCAGAGATTATACGATCTCTGCTGATGAAGCTTAACTTCTTACAAACATCTC 1020
Qy      1203  GATGCTTTTGGTATTATGACTATACCTTAACTTGAAGATGAGGATTAATCAGATGAC 1262
Db      1021  GATGCTTTTGGTATTATGACTATACCTTAACTTGAAGATGAGGATTAATCAGATGAC 1080
Qy      1263  ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAAATGCTAGCTAT 1322
Db      1081  ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAAATGCTAGCTAT 1140
Qy      1323  CATTTAGCCGGTGGT 1337
Db      1141  CATTTAGCCCTATGAT 1155

RESULT 3
LOCUS      I05204
DEFINITION Sequence 5 from Patent EP 0248227.
ACCESSION  I05204
VERSION     I05204.1 GI:591209
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE   1 (bases 1 to 1401)
AUTHORS     Hagenson,M.J. and Stroman,D.W.
TITLE       Yeast production of streptokinase
JOURNAL     Patent: EP 0248227-A1 5 09-DEC-1987;
FEATURES
            Location/Qualifiers
            source
              1..1401
              /organism="unknown"
              /mol_type="unassigned DNA"

ORIGIN
Query Match      69.2%; Score 1148.6; DB 6; Length 1401;
Best Local Similarity 99.2%; Pred. No. 1.9e-291;
Matches 1154; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy      175  ATACCATGATTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAACAGCCCAAT 234
Db      2  ATTCCATGATTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAACAGCCCAAT 61
Qy      235  TGGTTGTTAGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTT 294
Db      62  TAGTGTGTTAGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTT 121
Qy      295  TTGAAATCGATCTTAACATCAGACCTGCTCATGGAGGAAGACAGACCAAGGCTTAAGTC 354
Db      122  TTGAAATCGATCTTAACATCAGACCTGCTCATGGAGGAAGACAGACCAAGGCTTAAGTC 181
Qy      355  CAAATCAAAACCATTTGCTACTGTAGTAGTGGCCGCGATGTCACATAAACTTGAGAAAGCTG 414
Db      182  CAAATCAAAACCATTTGCTACTGTAGTAGTGGCCGCGATGTCACATAAACTTGAGAAAGCTG 241
Qy      415  ACTTACTAAAGGCTATTCAAGAACAAATTGATCGCTAACCGTCCACAGTAAACGAGCTACT 474
Db      242  ACTTACTAAAGGCTATTCAAGAACAAATTGATCGCTAACCGTCCACAGTAAACGAGCTACT 301
Qy      475  TTGAGTCAATTGATTGTCAGCGATGCAACCATTAATCTGATCGAAACGGCAAGGCTTACT 534
Db      302  TTGAGTCAATTGATTGTCAGCGATGCAACCATTAATCTGATCGAAACGGCAAGGCTTACT 361
Qy      535  TTGCTGACAAAGATGGTTCGGTAACTTCGGACCCCAACTGTCCAAGATTTTGTCTAA 594
```

```
Db      362  TTGCTGACAAAGATGGTTTGGTAACTTGGTCCGACCAACCTGTCCTCAAGAAATTTTGTCTAA 421
Qy      595  GCGGACATGTGCGGTTAGACCATATAAGAAAAACCAATACAAACCAACGGAATCTG 654
Db      422  GCGGACATGTGCGGTTAGACCATATAAGAAAAACCAATACAAACCAACGGAATCTG 481
Qy      655  TTGATGTGAATATATCTGTACAGATTTTACTCCCTTAAACCCCTGATGACGATTTTACAGCCAG 714
Db      482  TTGATGTGAATATATCTGTACAGATTTTACTCCCTTAAACCCCTGATGACGATTTTACAGCCAG 541
Qy      715  GTCTCAAGATATTAAGCTATTGAAAAACATAGCTATCGGTGACACCATCAATCTCAAG 774
Db      542  GTCTCAAGATATTAAGCTATTGAAAAACATAGCTATCGGTGACACCATCAATCTCAAG 601
Qy      775  AATTACTAGCTCAAGCACAAGACATTTTAAACAAAAACCCAGGCTATACGATTTATG 834
Db      602  AATTACTAGCTCAAGCACAAGACATTTTAAACAAAAACCCAGGCTATACGATTTATG 661
Qy      835  AACGTGACTCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGG 894
Db      662  AACGTGACTCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGG 721
Qy      895  ATCAAGAGTTTACTTACCGTGTAAATTCGGGAACAAGCTTATAGGATCAATAAAAAAT 954
Db      722  ATCAAGAGTTTACTTACCGTGTAAATTCGGGAACAAGCTTATAGGATCAATAAAAAAT 781
Qy      955  CTGCTCTGAATGAAGAAATAAACACTGACCTGATCTCTGAGAAATATTACGTCCTTA 1014
Db      782  CTGCTCTGAATGAAGAAATAAACACTGACCTGATCTCTGAGAAATATTACGTCCTTA 841
Qy      1015  AAAAAGGGGAAAGCCGTATGATCCCTTTGATCGCAGTCACCTTGAACCTGTTCCACATCA 1074
Db      842  AAAAAGGGGAAAGCCGTATGATCCCTTTGATCGCAGTCACCTTGAACCTGTTCCACATCA 901
Qy      1075  AATACGTTGATGTCGATACCAACGAATGTCTAAAAGTGAGCAGCTCTTAAACAGTAGCG 1134
Db      902  AATACGTTGATGTCGATACCAACGAATGTCTAAAAGTGAGCAGCTCTTAAACAGTAGCG 961
Qy      1135  AACGTAAGTTAGCTTCAGAGATTATACGATCTCTGCTGATGAAGCTTAACTTCTTACA 1194
Db      962  AACGTAAGTTAGCTTCAGAGATTATACGATCTCTGCTGATGAAGCTTAACTTCTTACA 1021
Qy      1195  ACAATCTCGATGCTTTTGGTATTATGACTATACCTTAACTGGAAGCTAGAGGATATC 1254
Db      1022  ACAATCTCGATGCTTTTGGTATTATGACTATACCTTAACTGGAAGCTAGAGGATATC 1081
Qy      1255  ACGATGACACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAGGAGAGATG 1314
Db      1082  ACGATGACACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAGGAGAGATG 1141
Qy      1315  CTAGCTATCATTTTAGCCGGTGGT 1337
Db      1142  TCAGCTATCATTTAGCCTATGAT 1164

RESULT 4
STRSKC
LOCUS      STRSKC
DEFINITION Streptococcus equisimilis (H46A) streptokinase gene, complete cds.
ACCESSION  K02986
VERSION     K02986.1 GI:153808
KEYWORDS    streptokinase.
SOURCE      Streptococcus dysgalactiae subsp. equisimilis
            Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
            Streptococcus.
            1 (bases 1 to 2568)
REFERENCE   1 (bases 1 to 2568)
AUTHORS     Malke,H., Roe,B. and Ferretti,J.J.
TITLE       Nucleotide sequence of the streptokinase gene from Streptococcus
            equisimilis H46A
JOURNAL     Gene 34 (2-3), 357-362 (1985)
PUBMED      2389113
COMMENT     Original source text: S.equisimilis (strain H46A) DNA, clone pMF5.
```

Draft entry and hard copy of sequence for [1] kindly provided by J.J. Ferretti, 03-SEP-1985.
The -35 and -10 regions are located at positions 760-765 and 781-786 respectively and an SD sequence at 808-813. Downstream from the coding region inverted repeats (positions 2176-2190 and 2203-2217) are thought to function as transcription terminators. The nucleotide sequence of skc does not support the hypothesis that the gene has evolved by duplication and fusion, as suggested by internal two-fold AA homologies of its product.

FEATURES

Location/Qualifiers
1..2568
/organism="Streptococcus dysgalactiae subsp. equisimilis"
/mol_type="genomic DNA"
/db_xref="taxon:119602"
794..>2141
/product="skc mRNA"
819..2141
/note="prestreptokinase"
/codon_start=1
/transl_table=11
/protein_id="AAA26974.1"
/db_xref="GI:153809"
/translation="MKNYLSFGMFALLFALTFTGVNSVOAIGPEWLDRSPVNSQLVSVAGTVEGTMDISLKFEDLTSRPAHGKTEQGLSPKSPFADTSGAMSHKLEKADLLKAIQBLTANYHSNDYFVIDFASDATITDRNGKVPADKDGSVTLPTQVQEFLLSGHVRVPYKEPIQNAQSVDEVTQPLNPDODFRPLKDTLLKTLAIGDTITSGELLAQAQSLINKHPGYTIYERDSSIYVHNDIIFRILPMQEBYFVRVNRQAYRINKSGLNEIINTDLISKYYVVKGERKPYDPFRSHLKLFTIKYVDVDNELLKSEQLLTASERLDFRDLVDRKAKLYNLDLDAFGIMDYTLTKGVKEDHNDTNRITVTMGRPEGENASVHLAYDKDRYTBEEREVSYLYRTGTPIDPNPNDK"
819..896
/note="streptokinase signal peptide"
897..2138
/product="streptokinase"
5 bp upstream of PstI site.
Query Match 69.2%; Score 1148.6; DB 1; Length 2568;
Best Local Similarity 99.7%; Pred. No. 1.9e-291;
Matches 1151; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 183 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAACGCAATTTGGTTGTT 242
Db 897 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAACGCAATTTAGTTGTT 956
Qy 243 AGCGTTGCTGTAAGTGGTGGAGGAGCAATCAAGACATTAAGTCTTAAATTTTTTGAATC 302
Db 957 AGCGTTGCTGTAAGTGGTGGAGGAGCAATCAAGACATTAAGTCTTAAATTTTTTGAATC 1016
Qy 303 GATCTAAACATCAGACCTCTCATGGAGGAAGACAGACGAGGCTTAAGTCCAAATCA 362
Db 1017 GATCTAAACATCAGACCTCTCATGGAGGAAGACAGACGAGGCTTAAGTCCAAATCA 1076
Qy 363 AAACCATTTGCTAGTATGTCGCGATGTGTCAATAAACTTGAGAAAGCTGACTTACTA 422
Db 1077 AAACCATTTGCTAGTATGTCGCGATGTGTCAATAAACTTGAGAAAGCTGACTTACTA 1136
Qy 423 AAGGCTATTCAAGAACAAATTGATCGTAAACGTCACAGTAACGACGACTTCTTGAAGTC 482
Db 1137 AAGGCTATTCAAGAACAAATTGATCGTAAACGTCACAGTAACGACGACTTCTTGAAGTC 1196
Qy 483 ATTGATTTTGCAGCGATGCAACCATTAATCTGATCGAAACGGCAGGCTTACTTTGCTGAC 542
Db 1197 ATTGATTTTGCAGCGATGCAACCATTAATCTGATCGAAACGGCAGGCTTACTTTGCTGAC 1256
Qy 543 AAAGATGTTTCCGTAACTTTCGCGACCCCACTGTCCAAAGAAATTTTGTCTAAGCGGCACAT 602
Db 1257 AAAGATGTTTCCGTAACTTTCGCGACCCCACTGTCCAAAGAAATTTTGTCTAAGCGGCAT 1316
Qy 603 GTGCGCGTTAGACCATATATAAGAAAAACCAATACAAAAAACGAGCAAAATCTGTTGATGTC 662
Db 1317 GTGCGCGTTAGACCATATATAAGAAAAACCAATACAAAAAACGAGCAAAATCTGTTGATGTC 1376
Qy 663 GAATATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGGCTCTCAA 722

Db 1377 GAATATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGGCTCTCAA 1436
Qy 723 GATACTAAGCTATTGAACAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTA 782
Db 1437 GATACTAAGCTATTGAACAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTA 1496
Qy 783 GCTCAAGCACAAGACATTTTAAACAAAAACACCCAGGCTATACGATTTTATGAACGTCAC 842
Db 1497 GCTCAAGCACAAGACATTTTAAACAAAAACACCCAGGCTATACGATTTTATGAACGTCAC 1556
Qy 843 TCCTCAATCGTCTACTCATGACAATGACATTTTCGGTACGATTTTACCAATGGATCAAGAG 902
Db 1557 TCCTCAATCGTCTACTCATGACAATGACATTTTCGGTACGATTTTACCAATGGATCAAGAG 1616
Qy 903 TTTTACTTACCGTGTAAATAATCGGGAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 962
Db 1617 TTTTACTTACCGTGTAAATAATCGGGAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 1676
Qy 963 AATGAAGAAATAAACACACCTGACCTCTCTGAGAAATATTACGTCCTTAAAAAAGGG 1022
Db 1677 AATGAAGAAATAAACACACCTGACCTCTCTGAGAAATATTACGTCCTTAAAAAAGGG 1736
Qy 1023 GAAAAGCCGTATGATCCCTTTTGATCGCAGTCACCTTGAACCTGTTTCAACATCAATACGTT 1082
Db 1737 GAAAAGCCGTATGATCCCTTTTGATCGCAGTCACCTTGAACCTGTTTCAACATCAATACGTT 1796
Qy 1083 GATGTCGATACCAACGAATTTGCTTAAAGTGAGCAGCTCTTAAACAGCTAGCGAAGCTAAC 1142
Db 1797 GATGTCGATACCAACGAATTTGCTTAAAGTGAGCAGCTCTTAAACAGCTAGCGAAGCTAAC 1856
Qy 1143 TTAGACTTCAGAGATTTTATACGATCCTCGTGATAAGGCTAACTTCTTCAACAATCTC 1202
Db 1857 TTAGACTTCAGAGATTTTATACGATCCTCGTGATAAGGCTAACTTCTTCAACAATCTC 1916
Qy 1203 GATGCTTTTGGTATTATGAGTATACCTTAACTTGAAGAGTAGAGGATAATCAGATGAC 1262
Db 1917 GATGCTTTTGGTATTATGAGTATACCTTAACTTGAAGAGTAGAGGATAATCAGATGAC 1976
Qy 1263 ACCAACCGTATACAAACCGTTTATATGGCAAGCAGCCGAGAGAGAGATGCTAGCTAT 1322
Db 1977 ACCAACCGTATACAAACCGTTTATATGGCAAGCAGCCGAGAGAGAGATGCTAGCTAT 2036
Qy 1323 CATTTAGCCGCGT 1337
Db 2037 CATTTAGCCGCTATGAT 2051
RESULT 5
A04926
LOCUS A04926
DEFINITION S.equisimilis skc gene for streptokinase.
ACCESSION A04926
VERSION A04926.1 GI:412219
KEYWORDS streptokinase.
SOURCE Streptococcus dysgalactiae subsp. equisimilis
ORGANISM Streptococcus dysgalactiae subsp. equisimilis
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.
REFERENCE 1 (bases 1 to 2568)
AUTHORS Ferretti, J.J. and Malke, H.
TITLE Streptokinase-coding recombinant vectors
JOURNAL Patent: EP 0151337-A 1 14-AUG-1985;
THE BOARD OF REGENTS OF THE UNIVERSITY OF OKLAHOMA
FEATURES
source Location/Qualifiers
1..2568
/organism="Streptococcus dysgalactiae subsp. equisimilis"
/mol_type="unassigned DNA"
/sub_species="equisimilis"
/db_xref="taxon:119602"

Query Match 69.2%; Score 1148.6; DB 6; Length 2568;

Best Local Similarity 99.7%; Pred. No. 1.9e-291;
Matches 1151; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 183 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCAATTTGGTGT 242
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 QY 897 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCAATTTGGTGT 956
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 QY 243 AGCGTTGCTGGTACTGTTGAGGGGAGCAATCAAGACATTAAGTCTTAAATTTTTTGAATC 302
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 QY 957 AGCGTTGCTGGTACTGTTGAGGGGAGCAATCAAGACATTAAGTCTTAAATTTTTTGAATC 1016
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 QY 303 GATCTAACATCAGCACTGCTCATGAGGAGAAAGACAGACAGCAAGCTTAAAGTCCAAATCA 362
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 QY 1017 GATCTAACATCAGCACTGCTCATGAGGAGAAAGACAGACAGCAAGCTTAAAGTCCAAATCA 1076
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 QY 363 AAACCAATTTGCTACTGATAGTGGCGGATGTCATATAAATTGAGAAAGCTGACTTACTA 422
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 QY 1077 AAACCAATTTGCTACTGATAGTGGCGGATGTCATATAAATTGAGAAAGCTGACTTACTA 1136
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 QY 423 AAGGCTATTCAAGAAACAATTTGATCGCTTAACGTCACAGTAACGACGACTACTTTGAGGTC 482
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 QY 1137 AAGGCTATTCAAGAAACAATTTGATCGCTTAACGTCACAGTAACGACGACTACTTTGAGGTC 1196
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 QY 483 ATTGATTTTGAAGCGATGCAACCATTAATCTGATCGAAACGCAAGGCTCTACTTTGCTGAC 542
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 QY 1197 ATTGATTTTGAAGCGATGCAACCATTAATCTGATCGAAACGCAAGGCTCTACTTTGCTGAC 1256
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 QY 543 AAGATGTTTCCGTACCTTCCGCGACCAACCTGTCGAAGAAATTTTCTTAAGCGGCAT 602
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 QY 1257 AAGATGTTTCCGTACCTTCCGCGACCAACCTGTCGAAGAAATTTTCTTAAGCGGCAT 1316
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 QY 603 GTGCGGTTAGACCATATAAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 662
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 QY 1317 GTGCGGTTAGACCATATAAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 1376
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 QY 663 GAATATAGTGTACGTTTACTCCCTTAAACCCCTGATGACGATTTACAGCAGGCTCTCAAA 722
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 QY 1377 GAATATAGTGTACGTTTACTCCCTTAAACCCCTGATGACGATTTACAGCAGGCTCTCAAA 1436
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 QY 723 GATCTAAGCTATTGAAACACTAGCTATCGGTGACCAATCATCATCTCAAGAAATTTACTA 782
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 QY 1437 GATCTAAGCTATTGAAACACTAGCTATCGGTGACCAATCATCATCTCAAGAAATTTACTA 1496
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 QY 783 GCTCAAGCAAAAGCAATTTTAAACAAAAACCCAGGCTATACGATTTATGACGTCAC 842
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 QY 1497 GCTCAAGCAAAAGCAATTTTAAACAAAAACCCAGGCTATACGATTTATGACGTCAC 1556
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 QY 843 TCTCAATGCTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 902
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 QY 1557 TCTCAATGCTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 1616
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 QY 903 TTTACTTACCGTGTAAAAATCGGGAACAAGCTTTATAGATCAATAAAAAATCTGGTCTG 962
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 QY 1617 TTTACTTACCGTGTAAAAATCGGGAACAAGCTTTATAGATCAATAAAAAATCTGGTCTG 1676
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 QY 963 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGCTTTAAAAAGGG 1022
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 QY 1677 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGCTTTAAAAAGGG 1736
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 QY 1023 GAAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAAACGTGTTACCATCAAAATAGTT 1082
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 QY 1737 GAAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAAACGTGTTACCATCAAAATAGTT 1796
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 QY 1083 GATGTCGATCAACAAAGAAATGCTAAAAAGTGAGCAGCTCTTAACAGCTAGCGAACGTAAC 1142
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 QY 1797 GATGTCGATCAACAAAGAAATGCTAAAAAGTGAGCAGCTCTTAACAGCTAGCGAACGTAAC 1856
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 QY 1143 TTAGACTTCAGAGATTTATACGATCTCTGATGATAAGCTTAACTACTCTACACAATCTC 1202
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 QY 1857 TTAGACTTCAGAGATTTATACGATCTCTGATGATAAGCTTAACTACTCTACACAATCTC 1916
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 QY 1203 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAGAGTAGAGGATATACGATGAC 1262
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 1917 GATGCTTTGGTATTATGAGCTATATACCTTAACTGAAAGTAGAGATAATCACGATGAC 1976
 QY 1263 ACCAACCGTATCATAAACCGTTTATATGCGCAAGCCAGCAGGAGAGAAATCTAGCTAT 1322
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1977 ACCAACCGTATCATAAACCGTTTATATGCGCAAGCCAGCAGGAGAGAAATCTAGCTAT 2036
 QY 1323 CATTTAGCCGCTGGT 1337
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 2037 CATTTAGCCCTATGAT 2051
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 RESULT 6
 SEDEXB
 LOCUS S.equisimilis dextb, abc, lrp, skc, rel genes and ORF1. 8931 bp DNA linear BCT 18-APR-2005
 DEFINITION X72832.1 GI:407876
 ACCESSION
 VERSION
 KEYWORDS abc gene; ABC transporter; dexB gene; dextran glucosidase; glucan 1,6-alpha-glucosidase; leucine rich protein; LRP gene; ORF1; rel gene; skc gene; streptokinase; stringent response-like protein.
 SOURCE Streptococcus dysgalactiae subsp. equisimilis
 ORGANISM Streptococcus dysgalactiae subsp. equisimilis
 Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Streptococcus.
 REFERENCE 1 (bases 3621 to 6190)
 AUTHORS Malke,H., Roe,B. and Ferretti,J.J.
 TITLE Nucleotide sequence of the streptokinase gene from Streptococcus equisimilis H46A
 JOURNAL Gene 34 (2-3), 357-362 (1985)
 PUBMED 298113
 REFERENCE 2 (bases 1 to 4188; 5790 to 8931)
 AUTHORS Mechold,U., Steiner,K., Vettermann,S. and Malke,H.
 TITLE Genetic organization of the streptokinase region of the Streptococcus equisimilis H46A chromosome
 JOURNAL Mol. Gen. Genet. 241 (1-2), 129-140 (1993)
 PUBMED 8232196
 REFERENCE 3 (bases 1 to 8931)
 AUTHORS Malke,H.
 TITLE Direct Submission
 JOURNAL Submitted (05-MAY-1993) H. Malke, Institute for Molecular Biology, Jena University, Winzerlaer Str 10, 07708 Jena, FRG
 COMMENT Related sequences: K02986, M19346, X13399 & X13400.
 FEATURES
 source
 1..8931
 /organism="Streptococcus dysgalactiae subsp. equisimilis"
 /mol_type="genomic DNA"
 /strain="H46A"
 /isolate="human group C strain"
 /sub_species="equisimilis"
 /db_xref="taxon:119602"
 /chromosome="streptokinase region"
 /clone_lib="E.coli plasmid library containing subfragments of the submitted seq:PSHD14/16, pSPV19, pSH2, pMP1, pCWP73, pRH10, pWX4"
 complement(89..1761)
 /gene="dexB"
 complement(89..115)
 /gene="dexB"
 /note="hairpin loop"
 complement(136..1749)
 /gene="dexB"
 /EC number="3.2.1.70"
 /codon_start=1
 /transl_table=11
 /product="glucan 1,6-alpha-glucosidase"
 /protein_id="CA51348.1"
 /db_xref="GI:407876"
 /db_xref="InterPro:IPR006047"
 /db_xref="InterPro:IPR006589"
 /db_xref="UniProt/Swiss-Prot:O59905"
 /translation="MQQVHHKATIIQIYPRSPKDTSGNGIGDLKIGTSDLYLQKLG ITAIWLSPVYQSPMDNDGYDISDYEAIEAFVGNMDDMLLAAANERGIKIIMDLVNN"
 gene
 terminator
 CDS

```

HTSDHAWFVEARENPNPSPERYIWRDEPNNLMSIFSGSAWELDEASQGYLHLFSK
KQPDNLNENAHVRQKIYDMWNFWIAKGIQGRMDVLDLIGKI PDSEITGNGRLHDYL
KEMNOATFGNHDVTGTWGAETARQYSRPENKELSMVFQFHVGLQHKNAKPM
DYAELDVPAKXTIIFSKWQTELKLGWNSLFWNNHDLPRVLSIWGNSDIYREKSACA
LAILLHMGRTIYIYQGEISGMTNYPFKDLTEVDDIESLNRYAKAEMNGVPAARVMS
TRKGRDNARTPMQSKDTHAGFSEAQETLVPVNPYQEIINADALANODSIPTTYQQ
LIALRKQDWLEADYHLLPTADKVFAYQRFGEETYVIVVNSDQEQFPAKDLAGAE
VVTNTVDVKLETHLQPDAPCVKLSV"
RBS complement(1757..1761)
gene /gene="dexB"
complement(1780..3051)
terminator /gene="abc"
complement(1780..1810)
misc_feature /gene="abc"
note="hairpin loop"
complement(2482..2499)
RBS /note="Walker motif B"
complement(2633..2965)
CDS /gene="abc"
/codon_start=1
/transl_table=11
/product="ABC transporter"
/protein_id="CAA51349.1"
/db_xref="GI:600072"
/db_xref="GOA:O54086"
/db_xref="InterPro:IPR003439"
/db_xref="UniProt/TREMBL:Q54086"
/misc_feature /translation="MVLENLNIHYKKYPNTHYAVEDFDLDIKDKBFI VFGVSGGCK
SYTLLMAGLEISIGELKIDGVEVNDKSPKORDIAMVFQNYALYPHMTVYDNNAFGL
KLRYKKCR"
complement(2644)
misc_feature /gene="abc"
note="(41) frame shift mutation in H46A"
complement(2828..2854)
misc_feature /gene="abc"
note="Walker motif A"
complement(2973..2976)
RBS /gene="abc"
complement(3026..3031)
-10_signal /gene="abc"
note="alternative"
complement(3043..3048)
-10_signal /gene="abc"
complement(3050..3051)
misc_feature /note="TG motif"
complement(3244..4149)
gene /gene="lrp"
complement(3244..3261)
terminator /gene="lrp"
note="hairpin loop"
complement(3262..4110)
CDS /gene="lrp"
/codon_start=1
/transl_table=11
/product="leucine rich protein"
/protein_id="CAA51350.1"
/db_xref="GI:407878"
/db_xref="GOA:O54087"
/db_xref="InterPro:IPR002197"
/db_xref="InterPro:IPR008931"
/misc_feature /db_xref="UniProt/Swiss-Prot:Q54087"
/translation="MELKDYFPFMQVGPPLGDKWVSVKEGQGYVHFPKSKLSEKER
LILLEVGLQGYEVLQSPWQYLLDHQGNPOLPQFETSPFIYLNHQVLPADLVLLIQ
QMAGLEVLTLPSTTQAFLCQATSIKVLRSLEGLLPTLESDFGLALTFVGNWYQ
VAAGTLRECFESECQLLTAYLKQKSGKLLTPAEVLMWSILSHQSFPALTRQTHQLN
PQSDMAVYHALWSEHNLVQTAQRDIYIHRNSLQYKLDKFAQQSGHLKQLDLDLAFAY
LFLKY"
complement(3337..3423)
misc_feature /gene="lrp"
note="leucine zipper motif"
complement(4117..4120)
RBS

```

```

-10_signal /gene="lrp"
complement(4141..4146)
misc_feature /gene="lrp"
complement(4148..4149)
misc_feature /gene="lrp"
note="TG motif"
4172..4368
note="static DNA bending locus"
4392..5837
gene /gene="skc"
4392..4393
misc_feature /gene="skc"
note="TG motif"
4395..4400
-10_signal /gene="skc"
4401..4406
-10_signal /gene="skc"
note="alternative"
4428..4431
RBS /gene="skc"
4439..5761
CDS /gene="skc"
/codon_start=1
/transl_table=11
/product="streptokinase"
/protein_id="CAA51351.1"
/db_xref="GI:407879"
/db_xref="GOA:P00779"
/db_xref="InterPro:IPR004093"
/db_xref="InterPro:IPR008124"
/db_xref="UniProt/Swiss-Prot:P00779"
/translation="MKNYLSFGMPALLFALTFGTVNSVQALAGPEWLLDRSVNNSQL
VVSVAGVEGNIQDISLKFEIDLITSRPAHGKTEQGLSPKSKPEFATDSGAMSHKLEK
ADLKAIOEQLIANVHNSDDYFEVIDFASDATIIDRNGKVYFADKQGSVTILPTQVQE
FLLSGHRVVRPYKEKPIQNAQSVDSYTVQFTPLNDDDFRGLKDTLKLTKLAIGD
TITSOELLAAQASILNKNHPGYTIYERDSSIIVTHNDIIFRILPMQDQFTYRVKNREQ
AYRLNKKSLGNEENNTDLISEKYVLKKGKPYDPPDRSHLKLFTIKYVDVDTNELL
KSEQLLTASERNLDFRLYDPRDKAKLLYNNLDAGIMDYTLTGKVENHDDTNRILIT
VYMKRPEGENASVYHLAYDKDRYTEERREYVSLRYTGTPIPDNPNDK"
4439..4516
sig_peptide /gene="skc"
5796..5837
terminator /gene="skc"
CDS complement(5859..6302)
note="unnamed protein product; ORF1"
/codon_start=1
/transl_table=11
/protein_id="CAA51352.1"
/db_xref="GI:407880"
/db_xref="GOA:O54088"
/db_xref="InterPro:IPR003732"
/db_xref="UniProt/Swiss-Prot:Q54088"
/translation="MKLVLRQKEASVSDIGKIAGAINQGLLLVGVGPDAAEDLAY
AVRKIVNMRIFSDADGKMNQSIODIKGSIILVSQFTLYADTKGNRPAPTGAAPDMA
SQFYDRENEOLADFPVVERGVFGADMVQVSLINDGFTVILDTKCH"
6007..6009
conflict /citation=[1]
replace="gc"
6162..6164
conflict /citation=[1]
replace="tg"
RBS complement(6310..6313)
note="ORF1"
gene complement(6317..8575)
note="rel"
CDS complement(6317..8536)
note="rel"

```

Query Match 69.2%; Score 1148.6; DB 1; Length 8931;
 Best local Similarity 99.7%; Pred. No. 2.1e-291;
 Matches 1151; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 183 ATTCTGTGACCTGAGTGTCTGCTAGACCGTCTGTCACACAGCAATGTTGTT 242

Db 4517 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCCAATTAGTTGTT 4576
Qy 243 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGGAAATC 302
Db 4577 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGGAAATC 4636
Qy 303 GATCTAAATCAGCACTGCTCATGTGAGGAGAAAGACAGAGCAAGGCTTTAAGTCCAAAATCA 362
Db 4637 GATCTAAATCAGCACTGCTCATGTGAGGAGAAAGACAGAGCAAGGCTTTAAGTCCAAAATCA 4696
Qy 363 AAACCAATTTGCTACTGATAGTGGCGGATGTCAATAAATTGAGAAAGCTGACTTACTA 422
Db 4697 AAACCAATTTGCTACTGATAGTGGCGGATGTCAATAAATTGAGAAAGCTGACTTACTA 4756
Qy 423 AAGGCTATTCAAGAACAAATTTGATCGCTTAACGTCCACAGTAGAACGACTACTTTGAGGTC 482
Db 4757 AAGGCTATTCAAGAACAAATTTGATCGCTTAACGTCCACAGTAGAACGACTACTTTGAGGTC 4816
Qy 483 ATTGAATTTTGAAGCGATGCAACCAATTACTGATCGAAACGGCAAGGTCTACTTTGCTGAC 542
Db 4817 ATTGAATTTTGAAGCGATGCAACCAATTACTGATCGAAACGGCAAGGTCTACTTTGCTGAC 4876
Qy 543 AAAGATGGTTCGGTAACCTTTCGCCGACCAACCTGTCCAGAAATTTTGTCTAAGCGGACAT 602
Db 4877 AAAGATGGTTCGGTAACCTTTCGCCGACCAACCTGTCCAGAAATTTTGTCTAAGCGGACAT 4936
Qy 603 GTGCGGTTAGACCATATAAGAAACCAATCAACCAACCAAGCGAATCTGTTGATGTG 662
Db 4937 GTGCGGTTAGACCATATAAGAAACCAATCAACCAACCAAGCGAATCTGTTGATGTG 4996
Qy 663 GAATATCTGTGACAGTTTACTCCCTTAAACCCCTGATGACGATTTCAGACAGGTCCTCAA 722
Db 4997 GAATATCTGTGACAGTTTACTCCCTTAAACCCCTGATGACGATTTCAGACAGGTCCTCAA 5056
Qy 723 GATCTAAGCTATTGAAAACACTGATCGGTGACACCATCATCATCTCAAGAAATTACTA 782
Db 5057 GATCTAAGCTATTGAAAACACTGATCGGTGACACCATCATCATCTCAAGAAATTACTA 5116
Qy 783 GCTCAAGCAAAAGCAATTTAAACAAACCAACCCAGGCTATAGGATTTATGAACGTGAC 842
Db 5117 GCTCAAGCAAAAGCAATTTAAACAAACCAACCCAGGCTATAGGATTTATGAACGTGAC 5176
Qy 843 TCCTCAATCGTCACTCATGACATGACATTTTCGTACGATTTTACCAATGGATCAAGAG 902
Db 5177 TCCTCAATCGTCACTCATGACATGACATTTTCGTACGATTTTACCAATGGATCAAGAG 5236
Qy 903 TTTACTTACCGTGTAAAAATCGGGAAACAGCTTTATAGGATCAATAAAAAATCTGGTCTG 962
Db 5237 TTTACTTACCGTGTAAAAATCGGGAAACAGCTTTATAGGATCAATAAAAAATCTGGTCTG 5296
Qy 963 AATGAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1022
Db 5297 AATGAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 5356
Qy 1023 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCCACCAATACGTT 1082
Db 5357 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCCACCAATACGTT 5416
Qy 1083 GATGTCGATACCAACCAATTTGCTTAAAGATGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 1142
Db 5417 GATGTCGATACCAACCAATTTGCTTAAAGATGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 5476
Qy 1143 TPAGACTTCAGAGATTTATACGATCTCTCGTGATAAGGCTAAACTACTCTACAAACATCTC 1202
Db 5477 TPAGACTTCAGAGATTTATACGATCTCTCGTGATAAGGCTAAACTACTCTACAAACATCTC 5536
Qy 1203 GATGCTTTTGGTATTATGACTATATCTTAACTTGGAAAAAGTAGAGGATAATCAGATGAC 1262
Db 5537 GATGCTTTTGGTATTATGACTATATCTTAACTTGGAAAAAGTAGAGGATAATCAGATGAC 5596
Qy 1263 ACCAACCGTATCATACCGTTTATATGGGCAAGCGGACCCGCAAGGAGAGATGCTAGTAT 1322

Db 5597 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAGGAGAGATGCTAGCTAT 5656
Qy 1323 CATTTAGCCGCTGGT 1337
Db 5657 CATTTAGCCCTATGAT 5671
RESULT 7
LOCUS E00522 2568 bp DNA linear PAT 29-SEP-1997
DEFINITION DNA fragment of plasmid PMF1 into which DNA encoding streptokinase is inserted.
ACCESSION E00522
VERSION E00522.1 GI:2168801
KEYWORDS JP 1985237995-A/1.
SOURCE Streptococcus dysgalactiae subsp. equisimilis
ORGANISM Streptococcus dysgalactiae subsp. equisimilis
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Streptococcus.
REFERENCE 1 (bases 1 to 2568)
AUTHORS Jiyosefu,J.F. and Horusuto,M.
TITLE RECOMINATION VECTOR FOR PRODUCING STREPTOKINASE
JOURNAL Patent: JP 1985237995-A 1 26-NOV-1985;
PHILLIPS PETROLEUM CO
COMMENT OS Streptococcus equisimilis
PN JP 1985237995-A/1
PD 26-NOV-1985
PF 09-OCT-1984 JP 1984212403
PR 10-OCT-1983 DD 83 255523, 02-MAR-1984 US 84 585417 PT
PJ JIYOSFU JIEI FUERTSUTEI, HORUSUTO MARUKE
PC C12N15/00, C12N1/20, C12N9/70, (C12N15/00, C12R1:46), (C12N1/20, PC C12R1:19),
PC (C12N9/70, C12R1:19);
CC strandedness: Single;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: strain=Streptococcus equisimilis H46A; CC *source: library=streptococcus equisimilis H46 library; CC *source: clone=lambda L47 skc clone;
FH Key Location/Qualifiers
FT 5'UTR 1..818
FT sig_peptide 819..896 /product='streptokinase signal peptide' FT
FT mat_peptide 897..2138 /product='streptokinase'
FT CDS 819..2143 /product='streptokinase'
FT 3'UTR 2142..2569.
FEATURES
source Location/Qualifiers
1..2568
/organism="Streptococcus dysgalactiae subsp. equisimilis"
/mol_type="genomic DNA"
/db_xref="taxon:119602"
ORIGIN
Query Match 68.9%; Score 1143.8; DB 6; Length 2568;
Best Local Similarity 99.4%; Pred. No. 3.6e-290;
Matches 1148; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Qy 183 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAAATTGGTTGTT 242
Db 897 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAAATTAGTTGTT 956
Qy 243 AGCGTTCGTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGGAAATC 302
Db 957 AGCGTTCGTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGGAAATC 1016
Qy 303 GATCTAAATCAGCACTGCTCATGTGAGGAGAAAGACAGAGCAAGGCTTTAAGTCCAAAATCA 362
Db 1017 GATCTAAATCAGCACTGCTCATGTGAGGAGAAAGACAGAGCAAGGCTTTAAGTCCAAAATCA 1076

QY 363 AAACCAATTGCTACTGATAGTGGCGGATGTGCATATAAACTTGAGAAAGCTGACTTACTA 422
Db |||||
1077 AAACCAATTGCTACTGATAGTGGCGGATGTGCATATAAACTTGAGAAAGCTGACTTACTA 1136
QY 423 AAGGCTATTCAAGAACAAATTGATCGGTAAAGTCCACAGTAAACGACGACTACTTTGAGGTC 482
Db |||||
1137 AAGGCTATTCAAGAACAAATTGATCGGTAAAGTCCACAGTAAACGACGACTACTTTGAGGTC 1196
QY 483 ATTGATTTTGAAGGATCAACCAATTACTGTATCGAAACGGCAAGGTCTACTTTGCTGAC 542
Db |||||
1197 ATTGATTTTGAAGGATCAACCAATTACTGTATCGAAACGGCAAGGTCTACTTTGCTGAC 1256
QY 543 AAGATGTTTCCGTTACCTTGGCGACCAACCTGTCCAAAGAAATTTTGTCTAAACGGACAT 602
Db |||||
1257 AAGATGTTTCCGTTACCTTGGCGACCAACCTGTCCAAAGAAATTTTGTCTAAACGGACAT 1316
QY 603 GTGCGGTTAGACCATATAAAGAAACCAATAACAAACCAAGCAAACTGTGTGATGTG 662
Db |||||
1317 GTGCGGTTAGACCATATAAAGAAACCAATAACAAACCAAGCAAACTGTGTGATGTG 1376
QY 663 GAAATATCTGTACAGTTTACTCCCTTAAACCTGTATGACGATTTTCAACGAGTCTCAAA 722
Db |||||
1377 GAAATATCTGTACAGTTTACTCCCTTAAACCTGTATGACGATTTTCAACGAGTCTCAAA 1436
QY 723 GATCTAAGCTATTGAACACACTAGCTATCGGTGACACCATCACTCTCAAGAAATTACTA 782
Db |||||
1437 GATCTAAGCTATTGAACACACTAGCTATCGGTGACACCATCACTCTCAAGAAATTACTA 1496
QY 783 GCTCAAGCACAAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTTATGAACGTGAC 842
Db |||||
1497 GCTCAAGCACAAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTTATGAACGTGAC 1556
QY 843 TCCTCAATCGTCACTCATGACAAATGACATTTTCGGTACGATTTTACCAATGGATCAAGAG 902
Db |||||
1557 TCCTCAATCGTCACTCATGACAAATGACATTTTCGGTACGATTTTACCAATGGATCAAGAG 1616
QY 903 TTTACTACCGTGTAAACCTCGGAAACCAAGCTTATAGGATCAATAAATAAATCTGGTCTG 962
Db |||||
1617 TTTACTACCGTGTAAACCTCGGAAACCAAGCTTATAGGATCAATAAATAAATCTGGTCTG 1676
QY 963 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAATAAGGG 1022
Db |||||
1677 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAATAAGGG 1736
QY 1023 GAAAGCGGTATGATCCCTTTGATGCGAGTCACTTGAACACTGTTTCCATCAAAATAGGTT 1082
Db |||||
1737 GAAAGCGGTATGATCCCTTTGATGCGAGTCACTTGAACACTGTTTCCATCAAAATAGGTT 1796
QY 1083 GATGTCGATACCAACGAAATGCTTAAAGTGAAGCAGCTCTTAACAGCTAGCGAACGTAAC 1142
Db |||||
1797 GATGTCGATACCAACGAAATGCTTAAAGTGAAGCAGCTCTTAACAGCTAGCGAACGTAAC 1856
QY 1143 TTAGACTTTCAGAGATTATACGATCTCTGATGATAGGCTTAAAGTGAAGCAGCTCTTAACAGCTAGCGAACGTAAC 1202
Db |||||
1857 TTAGACTTTCAGAGATTATACGATCTCTGATGATAGGCTTAAAGTGAAGCAGCTCTTAACAGCTAGCGAACGTAAC 1916
QY 1203 GATGCTTTTGGTATTATGAGCTATACCTTAACTGGAAGAGTAGAGGATATCAAGTAC 1262
Db |||||
1917 GATGCTTTTGGTATTATGAGCTATACCTTAACTGGAAGAGTAGAGGATATCAAGTAC 1976
QY 1263 ACCAACCGTATCAACCGTTTATATGGGCAAGCGACCGGAGGAGATGCTAGCTAT 1322
Db |||||
1977 ACCAACCGTATCAACCGTTTATATGGGCAAGCGACCGGAGGAGATGCTAGCTAT 2036
QY 1323 CATTTAGCGGCGGT 1337
Db |||||
2037 CATTTAGCGCTATGAT 2051

RESULT 8
QY97820 7057 bp DNA linear PAT 20-APR-2004
LOCUS QY97820
DEFINITION Sequence 26 from Patent WO2004029256.

ACCESSION CQ797820
VERSION CQ797820.1 GI:46426093
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS selman-housein Sosa,G., aguiar Cabeza,E., gonz lez Quintero,A.D.
and ramos gonz Lez,O.
Vector for the production of transplastomic angiosperm plants
Patent: WO 2004029256-A 26 08-APR-2004;
CENTRO DE INGENIERIA GENETICA Y BIOTECNOLOGIA (CU)
FEATURES
source 1..7057
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Description of Artificial Sequence: Nucleotide
sequence of DNA fragment from the vector pVTPA-Estrep
between the rice atpB and tobacco rbcL borders."
ORIGIN
Query Match 68.4%; Score 1136.2; DB 6; Length 7057;
Best Local Similarity 98.5%; Pred. No. 3.8e-288;
Matches 1147; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
QY 173 ATATACCATGATTCGTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCCA 232
Db |||||
2013 ATAAACCATGATTCGTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCCA 2072
QY 233 ATTGGTTGTTAGCGTTCCTGCTACTGTTGAGGGGACGAATCAAGACATTTAGTCTTTAAATT 292
Db |||||
2073 ATTAGTTGTTAGCGTTCCTGCTACTGTTGAGGGGACGAATCAAGACATTTAGTCTTTAAATT 2132
QY 293 TTTTGAATCGATCTAAACATCACGACCTGCTCATGAGGAGAAAGACAGACGAAGCTTAAAG 352
Db |||||
2133 TTTTGAATTTGACCTTAACATCACGACCTGCTCATGAGGAGAAAGACAGACGAAGCTTAAAG 2192
QY 353 TCACAAATCAAAACCAATTTGCTACTGATAGTGGCGGATGTCAATAACTTTGAGAAAGC 412
Db |||||
2193 TCCAAATCAAAACCAATTTGCTACTGATAGTGGCGGATGTCCATAACTTTGAAAAAGC 2252
QY 413 TGACTTACTAAAGCGTATTCAAGAACAAATTTGATCGCTAAACGTCACACAGTACGACGACTA 472
Db |||||
2253 TGACTTACTAAAGCGTATTCAAGAACAAATTTGATCGCTAAACGTCACACAGTACGACGACTA 2312
QY 473 CTTTGAAGTCATTTGATTTTGCACGCGATGCAACCATTTACTGATCGAAACGGCAAGGCTCTA 532
Db |||||
2313 CTTTGAAGTCATTTGATTTTGCACGCGATGCAACCATTTACTGATCGAAACGGCAAGGCTCTA 2372
QY 533 CTTTGTGACAAAGATGGTTTCGGTTAACTTTGCGGACCCCAACCTGTCCAAGAAATTTTGTCT 592
Db |||||
2373 CTTTGTGACAAAGATGGTTTCGGTTAACTTTGCGGACCCCAACCTGTCCAAGAAATTTTGTCT 2432
QY 593 AAGCGGACATGTCGCGCTTAGACCATATAAAGAAACCAATCAAAACCAAGCGAAATC 652
Db |||||
2433 AAGCGGACATGTCGCGCTTAGACCATATAAAGAAACCAATCAAAACCAAGCGAAATC 2492
QY 653 TGTGTGATGCGAATATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTCAGAC 712
Db |||||
2493 TGTGTGATGCGAATATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTCAGAC 2552
QY 713 AGGTCTCAAAAGATCTAAGCTATTGAAAAACA CTAGCTATCGGTGACACCATCAATCTCA 772
Db |||||
2553 AGGTCTCAAAAGATCTAAGCTATTGAAAAACA CTAGCTATCGGTGACACCATCAATCTCA 2612
QY 773 AGAATTAAGTCTCAAGCACAAAGCATTTTAAACAAAAACCCACCGGCTATACGATTTA 832
Db |||||
2613 AGAATTAAGTCTCAAGCACAAAGCATTTTAAACAAAAACCCACCGGCTATACGATTTA 2672
QY 833 TGAACGTGACTCTCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACC AAT 892
Db |||||
2673 TGAACGTGACTCTCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACC AAT 2732

QY 893 GGATCAAGAGTTTACTTACCGTGTAAAAATCGGGAAACAAGCTTATAGGATCAATAAAAA 952
Db 2733 GGATCAAGAGTTTACTTACCATGTGCAAAAATCGGAAACAAGCTTATGAGATCAATAAAAA 2792
QY 953 ATCTGCTCTGAATGAAGAAATAACAACACTGACCTGATCTCTGAGAAATATTACGTCCT 1012
Db 2793 ATCTGCTCTGAATGAAGAAATAACAACACTGACCTGATCTCTGAGAAATATTACGTCCT 2852
QY 1013 TAAAAAGGGGAAAAAGCCGTATGATCCTTTGATCGCAGTCACCTTGAACCTGTTTACCACAT 1072
Db 2853 TAAAAAGGGGAAAAAGCCGTATGATCCTTTGATCGCAGTCACCTTGAACCTGTTTACCACAT 2912
QY 1073 CAAATACGTTGATGTGCGATACCAACGAATGCTTAAAAAGTGAGCAGCTCTTAAACAGCTAG 1132
Db 2913 CAAATACGTTGATGTCAACACCAACGAATGCTTAAAAAGTGAGCAGCTCTTAAACAGCTAG 2972
QY 1133 CGAACGTAACCTTAGACTTCAGAGATTATACGATTCCTCGTGATAAGGCTTAACTACTCTTA 1192
Db 2973 CGAACGTAACCTTAGACTTCAGAGATTATACGATTCCTCGTGATAAGGCTTAACTACTCTTA 3032
QY 1193 CAACAATCTCGATGCTTTTGGTATTATGGAATTAACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAA 1252
Db 3033 CAACAATCTCGATGCTTTTGGTATTATGGAATTAACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAA 3092
QY 1253 TCACGATGACACCAACCGCTATCAACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAA 1312
Db 3093 TCACGATGACACCAACCGCTATCAACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAA 3152
QY 1313 TGCTAGCTATCATTTAGCCGGTGGT 1337
Db 3153 TGCTAGCTATCATTTAGCCCTATGAT 3177

RESULT 9
AR068768
LOCUS
DEFINITION
Sequence 19 from patent US 5854049.
ACCESSION
AR068768
VERSION
AR068768.1 GI:6000975
KEYWORDS
Unknown.
SOURCE
Unknown.
ORGANISM
Unclassified.
REFERENCE
1 (bases 1 to 2566)
AUTHORS
Reed,G.L.
TITLE
Plasmin-resistant streptokinase
JOURNAL
Patent: US 5854049-A 19 29-DEC-1998;
FEATURES
Location/Qualifiers
source
1..2566
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 68.3%; Score 1134.6; DB 6; Length 2566;
Best Local Similarity 99.5%; Pred. No. 9.5e-288;
Matches 1149; Conservative 0; Mismatches 4; Indels 2; Gaps 1;

QY 183 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAACAGCCCAATTTGGTTGTT 242
Db 897 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAACAGCCCAATTTAGTTGTT 956
QY 243 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAATC 302
Db 957 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAATC 1016
QY 303 GATCTAAACATCAGACCTGCTCATGAGGAGAGAGACAGACGACGCTTAACTCAAAATCA 362
Db 1017 GATCTAAACATCAGACCTGCTCAT--AGGAAAGACAGACGACGCTTAACTCAAAATCA 1074
QY 363 AAACCAATTTGCTACTGATAGTGGCGGATGTACATAAACTTGAAGAAAGCTGACTTACTA 422
Db 1075 AAACCAATTTGCTACTGATAGTGGCGGATGTACATAAACTTGAAGAAAGCTGACTTACTA 1134

QY 423 AAGGCTATTTCAAGAACAAATTGATCGCTAAACGTCCACAGTAACGACGACTACTTTTCAGGTC 482
Db 1135 AAGGCTATTTCAAGAACAAATTGATCGCTAAACGTCCACAGTAACGACGACTACTTTTCAGGTC 1194
QY 483 ATTGATTTTGCAGGATGCAACCAATTAATCTGATCGAAACGGCAAGGTCTACTTTCTGAC 542
Db 1195 ATTGATTTTGCAGGATGCAACCAATTAATCTGATCGAAACGGCAAGGTCTACTTTCTGAC 1254
QY 543 AAAGATGTTGCGGTAAACCTTGCAGCCCAACCAATTAATCTGATCGAAACGGCAAGGTCTACTTTCTGAC 602
Db 1255 AAAGATGTTGCGGTAAACCTTGCAGCCCAACCAATTAATCTGATCGAAACGGCAAGGTCTACTTTCTGAC 1314
QY 603 GTGCGCGTTAGACCATATAAAGAAAAACCAATTAATCTGATCGAAACGGCAAGGTCTACTTTCTGAC 662
Db 1315 GTGCGCGTTAGACCATATAAAGAAAAACCAATTAATCTGATCGAAACGGCAAGGTCTACTTTCTGAC 1374
QY 663 GAATATATCTGATGAGTTTACTCCCTTAAACCTGATGACGATTTTTCAGACCGAGGTCTCAAA 722
Db 1375 GAATATATCTGATGAGTTTACTCCCTTAAACCTGATGACGATTTTTCAGACCGAGGTCTCAAA 1434
QY 723 GATATTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCAATCTCAAGAAATTACTA 782
Db 1435 GATATTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCAATCTCAAGAAATTACTA 1494
QY 783 GCTCAAGCACAAAGCATTTTAAACAAAAACCAACCCAGGCTATACGATTTATGAACGTGAC 842
Db 1495 GCTCAAGCACAAAGCATTTTAAACAAAAACCAACCCAGGCTATACGATTTATGAACGTGAC 1554
QY 843 TCCTCAATCGTCACTCATGACAAATGACATTTTTCGGTACGATTTTACCAATGGATCAAGAG 902
Db 1555 TCCTCAATCGTCACTCATGACAAATGACATTTTTCGGTACGATTTTACCAATGGATCAAGAG 1614
QY 903 TTACTTACCGTGTAAAAATCGGGAAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 962
Db 1615 TTACTTACCGTGTAAAAATCGGGAAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 1674
QY 963 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAGCTCTTAAAAAAGG 1022
Db 1675 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAGCTCTTAAAAAAGG 1734
QY 1023 GAAAAGCGGTATGATCCTTTGATCGAGTCACCTTGAACCTGTTCAACCATCAAAATACGTT 1082
Db 1735 GAAAAGCGGTATGATCCTTTGATCGAGTCACCTTGAACCTGTTCAACCATCAAAATACGTT 1794
QY 1083 GATGTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTTAACAGCTTAGCGAAACGTAAC 1142
Db 1795 GATGTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTTAACAGCTTAGCGAAACGTAAC 1854
QY 1143 TTAGACTTCAGAGATTTTATACGATCCTCGTGATAGGCTAACTACTCTACAAATCTC 1202
Db 1855 TTAGACTTCAGAGATTTTATACGATCCTCGTGATAGGCTAACTACTCTACAAATCTC 1914
QY 1203 GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAGAGTAGAGGATAATCAGATGAC 1262
Db 1915 GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAGAGTAGAGGATAATCAGATGAC 1974
QY 1263 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAAATGCTAGCTAT 1322
Db 1975 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAAATGCTAGCTAT 2034
QY 1323 CATTTAGCCGGTGGT 1337
Db 2035 CATTTAGCCCTATGAT 2049

RESULT 10
S46536
LOCUS
DEFINITION
SKC-2=streptokinase [Streptococcus equisimilis, group C, ATCC 9542, Genomic, 1245 nt].
ACCESSION
S46536
VERSION
S46536.1 GI:257196

KEYWORDS

SOURCE Streptococcus dysgalactiae subsp. equisimilis
ORGANISM Streptococcus dysgalactiae subsp. equisimilis
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.

REFERENCE

AUTHORS Estrada,M.P., Hernandez,L., Perez,A., Rodriguez,P., Serrano,R.,
Rubiera,R., Pedraza,A., Padron,G., Antuch,W., de la Fuente,J. et
al.

TITLE High level expression of streptokinase in Escherichia coli

JOURNAL Biotechnology (N.Y.) 10 (10), 1138-1142 (1992)

PUBMED 1368792

REMARK GenBank staff at the National Library of Medicine created this
entry [NCBI gibbsq 115306] from the original journal article.

FEATURES

source
1..1245
/organism="Streptococcus dysgalactiae subsp. equisimilis"
/mol_type="genomic DNA"
/db_xref="taxon:119602"
/note="group: C"

gene

1..1245

/gene="SKC-2"

1..1245

/gene="SKC-2"

/note="streptokinase; Mature protein lacks initial Met"

/codon_start=1

/product="SKC-2"

/protein_id="AAC60418.1"

/db_xref="GI:257197"

/translation="TAGPEWLLDRPSVNNLSQVSVAGTVGEGTNOISLKFPEDILTS
PRAHGKTEOGLSPKSPKPPATDSGAMPHKLEKADLLKAIOELIANVHNSDNDYFEVID
FASDATIIDRNGKVVFAKDGSVTLPTQVDFELLSCHVVRVPYKEKIQNAKSVDV
EYTVQFTFLNPDPPDFRPLGKDKLKTALGDTTISOELLAAQSLINKHPHYIYE
RDSSIVTHNDIFRITLPMQBFTHVNRQAYEINKSGLEENINNTDLISEKYIV
LKKGKPYDPFDRSHLPLTIKYVDVNTNELLSEQLLTASERNLDFRDLYPDPRXAK
LLYNLDAFGIMDYTLTGKVEDNHDHTNRITVYMGKREGENASVHLAYDKDRYTEE
EREVYSYLRYTGTPIPDNENDK"

ORIGIN

Query Match 67.9%; Score 1127.8; DB 1; Length 1245;
Best Local Similarity 98.5%; Pred. No. 5.7e-286;
Matches 1138; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 183 ATTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAGACCAATTTGGTTGT 242
Db 1 ATTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAGACCAATTTAGTTGT 60
Qy 243 AGCGTTGCTGTGCTGTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTGAAATC 302
Db 61 AGCGTTGCTGTGCTGTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTGAAAT 120
Qy 303 GATCTAAATCATCAGACCTGCTCATGGAGGAAAGACAGACAGGCTTAAAGTCCAAATCA 362
Db 121 GACCTAAATCATCAGACCTGCTCATGGAGGAAAGACAGACAGGCTTAAAGTCCAAATCA 180
Qy 363 AAACCAATTTGCTGATGATGAGTGGCGGATGTCATATAAAGTGAAGAGCTGACTACTA 422
Db 181 AAACCAATTTGCTGATGATGAGTGGCGGATGTCATATAAAGTGAAGAGCTGACTACTA 240
Qy 423 AAGGCTATTCAAGAACAAATTTGATCGGTAAAGTCAAGTAAAGACGACTACTTTGAGTTC 482
Db 241 AAGGCTATTCAAGAACAAATTTGATCGGTAAAGTCAAGTAAAGACGACTACTTTGAGTTC 300
Qy 483 ATTGATTTTGAAGCGATGCAACCAATTTACTGATCGAAACGGCAAGGCTCTACTTTGCTGAC 542
Db 301 ATTGATTTTGAAGCGATGCAACCAATTTACTGATCGAAACGGCAAGGCTCTACTTTGCTGAC 360
Qy 543 AAAGATGGTTCGGTAACTTGGCGGACCAACCTGTCCAGAAATTTTGTCTAAGCGGACAT 602
Db 361 AAAGATGGTTCGGTAACTTGGCGGACCAACCTGTCCAGAAATTTTGTCTAAGCGGACAT 420
Qy 603 GTGCGCGTTAGACCATATAAAGAAACCAATATACAAACCAAGCGAAATCTGTGTGATGTC 662

Db 421 GTCCGCGTTAGACCATATAAAGAAAAACCAATATACAAATCAAGCGAAATCTGTTGATGTG 480
Qy 663 GAATATACCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGGTCTCAAA 722
Db 481 GAATATACCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGGTCTCAAA 540
Qy 723 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCACTCATCATCTCAAGAAATTACTA 782
Db 541 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCACTCATCATCTCAAGAAATTACTA 600
Qy 783 GCTCAAGACCAAGCAATTTTAAACCAAAACCAACCCAGGCTATACGATTTATGAACGTGAC 842
Db 601 GCTCAAGACCAAGCAATTTTAAACCAAAACCAACCCAGGCTATACGATTTATGAACGTGAC 660
Qy 843 TCCTCAATCGTCACTCATGACCAATTTTCCGTACGATTTTCCGTACGATTTTACCAATGGATCAAGAG 902
Db 661 TCCTCAATCGTCACTCATGACCAATTTTCCGTACGATTTTCCGTACGATTTTACCAATGGATCAAGAG 720
Qy 903 TTTTACTTACCGTGTAAATAATCGGAAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 962
Db 721 TTTTACTTACCGTGTCAATAATCGGAAACAAGCTTTATGAGATCAATAAAAAATCTGGTCTG 780
Qy 963 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 1022
Db 781 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 840
Qy 1023 GAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTTCCACATCAATAGCTT 1082
Db 841 GAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTTCCACATCAATAGCTT 900
Qy 1083 GATGTGCTGATCAACCAAGAAATGCTTAAAGTGAGCAGCTCTTAAACAGCTTAGCGAACGTAAC 1142
Db 901 GATGTCAACCAACCAAGAAATGCTTAAAGTGAGCAGCTCTTAAACAGCTTAGCGAACGTAAC 960
Qy 1143 TTAGACTTTCAGAGATTTTATACGATCTCTGTAAGCTAAAGCTAAACTCTTACAAACAATCTC 1202
Db 961 TTAGACTTTCAGAGATTTTATACGATCTCTGTAAGCTAAAGCTAAACTCTTACAAACAATCTC 1020
Qy 1203 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAGAGTAGGAGGATATACAGATGAC 1262
Db 1021 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAGAGTAGGAGGATATACAGATGAC 1080
Qy 1263 ACCAAGCGTATCATACCGTTTATATGGCAAGCGCCGCAAGGAGAGATGCTAGCTAT 1322
Db 1081 ACCAAGCGTATCATACCGTTTATATGGCAAGCGCCGCAAGGAGAGATGCTAGCTAT 1140
Qy 1323 CATTTAGCCGTTGGT 1337
Db 1141 CATTTAGCCCTATGAT 1155

RESULT 11

A20015
LOCUS SEQ ID NO: 23; Nucleotide sequence for methionol-streptokinase
DEFINITION fusion protein.
ACCESSION A20015
VERSION A20015.1 GI:1247848
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE synthetic sequences.
1 (bases 1 to 1257)
AUTHORS PROTEINS AND NUCLEIC ACIDS
TITLE Patent: WO 9109125-A 23 27-JUN-1991;
JOURNAL Location/Qualifiers
FEATURES
source
1..1257
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
4..1251
/codon_start=1
CDS

```
/transl_table=11
/product="methionyl-streptokinase fusion protein"
/protein_id="CAA01486.1"
/db_xref="GI:1247849"
/translation="MTAGPEWLLDRSPVNNNSOLVSVAGTVEGTNQDISLKFPEDIDT
SPAPGKTEQGLSPKSKPENTDSGAMPKLEKADLLKAQELIANVHNDIDFEVI
DFASADITDRNGKVTFADKDSVTLPTQVBFLLSGHVRVPRYPKPKIONAKSVYD
VEYTFQFTPLNPDPPFGLKDTLLKTLAIGDTITSQELLAQAQSLINAKHPGYTIY
ERDSSTVTHDNDIFRILPMDQBFYVHVKNREQAYEINKKSLNEEINNTDLISEKY
VLKGEKPYDPPDRSHLKLFTIKYVDVNTNELLKSEOLLTKSERNLDPRLDYPDRKA
KLLNNLDAFGIMDYTLTGKVENHDDTNRIITVYMKRPEGENASHLAYDKDRYTE
EBREYVSYLRYGTGPIPDNDK"
```

ORIGIN

```
Query Match 67.9%; Score 1127.6; DB 6; Length 1257;
Best Local Similarity 98.4%; Pred. No. 6.5e-286;
Matches 1139; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 180 ATGATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCCAATTGGTT 239
Db 4 ATGATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCCAATTAGTT 63

Qy 240 GTTAGCGTTGCTGGTACTGTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGA 299
Db 64 GTTAGCGTTGCTGGTACTGTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGA 123

Qy 300 ATCGATCTAACTACGACCTGCTCATGAGGAAAGACAGACAGAGGCTTAAAGTCCAAA 359
Db 124 ATTGACCTTAACATCAGACCTGCTCATGAGGAAAGACAGACAGAGGCTTAAAGTCCAAA 183

Qy 360 TCAAAACCAATTTGCTACTGATAGTGGCGGATGTCATATAAATTTGAGAAAGCTGACTTA 419
Db 184 TCAAAACCAATTTGCTACTGATAGTGGCGGATGTCACATATAAATTTGAGAAAGCTGACTTA 243

Qy 420 CTAAGGCTATTCAAGAACAAATTGATCGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 479
Db 244 CTAAGGCTATTCAAGAACAAATTGATCGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 303

Qy 480 GTCAATTGATTTGCAAGCGATCAACCATTTACTGATCGAAACCGCAGGCTCTACTTTGCT 539
Db 304 GTCAATTGATTTGCAAGCGATCAACCATTTACTGATCGAAACCGCAGGCTCTACTTTGCT 363

Qy 540 GACAAAGATGGTTGCGTAACTTGGCCGACCAACCTGTCCAGAAATTTTGTAAAGCGGA 599
Db 364 GACAAAGATGGTTGCGTAACTTGGCCGACCAACCTGTCCAGAAATTTTGTAAAGCGGA 423

Qy 600 CATGTGCGGTTAGACCATATAAAGAAACCAATAAACAACCAAGCGAAATCTGTGTAT 659
Db 424 CATGTGCGGTTAGACCATATAAAGAAACCAATAAACAACCAAGCGAAATCTGTGTAT 483

Qy 660 GTGGAATATACGTACAGTTTACTCCCTTAAACCTGTATGACGATTTCAAGACAGGCTCT 719
Db 484 GTGGAATATACGTACAGTTTACTCCCTTAAACCTGTATGACGATTTCAAGACAGGCTCT 543

Qy 720 AAAGATACCTAAGCTATTGAAACACATAGTATCGGTGACACCATCATCTCAAGAAATTA 779
Db 544 AAAGATACCTAAGCTATTGAAACACATAGTATCGGTGACACCATCATCTCAAGAAATTA 603

Qy 780 CTAGCTCAAGCAAAAGCAATTTTAAACAAAACCAACCGGCTATACGATTTATGAACGT 839
Db 604 CTAGCTCAAGCAAAAGCAATTTTAAACAAAACCAACCGGCTATACGATTTATGAACGT 663

Qy 840 GACTCCTCAATCGTCACTCATGACAAATGACATTTTCGGTACGATTTTACCAATGGATCAA 899
Db 664 GACTCCTCAATCGTCACTCATGACAAATGACATTTTCGGTACGATTTTACCAATGGATCAA 723

Qy 900 GAGTTTACTTACGGTGTAAATTCGGGACACAGCTTATAGGATCAATAAAATCTGGT 959
Db 724 GAGTTTACTTACGGTGTAAATTCGGGACACAGCTTATAGGATCAATAAAATCTGGT 783

Qy 960 CTGAATGAAGAAATAAACAACCTGACCTGATCTCTGAGAAATATTACGCTCTTAAAAA 1019
Db 784 CTGAATGAAGAAATAAACAACCTGACCTGATCTCTGAGAAATATTACGCTCTTAAAAA 843
```

```
Qy 1020 GGGGAAAGCCGTATGATCCCTTTTATCCAGTCACTTGAACACTGTTCAACATCAATAC 1079
Db 844 GGGGAAAGCCGTATGATCCCTTTTATCCAGTCACTTGAACACTGTTCAACATCAATAC 903

Qy 1080 GTTGATGTCGATACCAACGAATTTGCTAAAGTGAAGAGCTTTAAACAGCTAGCGAAGT 1139
Db 904 GTTGATGTCGATACCAACGAATTTGCTAAAGTGAAGAGCTTTAAACAGCTAGCGAAGT 963

Qy 1140 AACTTAGACTTTCAGAGATTTTATAGATCCTCGTGAAGGCTTAACTACTCTACAACT 1199
Db 964 AACTTAGACTTTCAGAGATTTTATAGATCCTCGTGAAGGCTTAACTACTCTACAACT 1023

Qy 1200 CTCGATGCTTTTGGTATTATGCACTATACCTTTAACTGAAAGTAGAGATAATCACCAT 1259
Db 1024 CTCGATGCTTTTGGTATTATGCACTATACCTTTAACTGAAAGTAGAGATAATCACCAT 1083

Qy 1260 GACACCAACCGTATCATAACCGTTTATATGGGCAAGCGACCGAGAGAGAAATGCTAGC 1319
Db 1084 GACACCAACCGTATCATAACCGTTTATATGGGCAAGCGACCGAGAGAGAAATGCTAGC 1143

Qy 1320 TATCATTTAGCCGCTGGT 1337
Db 1144 TATCATTTAGCCCTATGAT 1161

RESULT 12
I13203 LOCUS I13203 1257 bp DNA linear PAT 26-JUL-1995
DEFINITION Sequence 25 from patent US 5434073.
ACCESSION I13203
VERSION I13203.1 GI:910551
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1257)
AUTHORS Dawson, K., Hunter, M. G. and Czaplewski, L. G.
TITLE Fibrinolytic and anti-thrombotic cleavable dimers
JOURNAL Patent: US 5434073-A 25 18-JUL-1995;
FEATURES Location/Qualifiers
source 1..1257
/organism="unknown"
/mol_type="unassigned DNA"
```

```
Query Match 67.9%; Score 1127.6; DB 6; Length 1257;
Best Local Similarity 98.4%; Pred. No. 6.5e-286;
Matches 1139; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 180 ATGATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCCAATTGGTT 239
Db 4 ATGATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCCAATTAGTT 63

Qy 240 GTTAGCGTTGCTGGTACTGTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGA 299
Db 64 GTTAGCGTTGCTGGTACTGTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGA 123

Qy 300 ATCGATCTAACTACGACCTGCTCATGAGGAAAGACAGACAGAGGCTTAAAGTCCAAA 359
Db 124 ATTGACCTTAACATCAGACCTGCTCATGAGGAAAGACAGACAGAGGCTTAAAGTCCAAA 183

Qy 360 TCAAAACCAATTTGCTACTGATAGTGGCGGATGTCATATAAATTTGAGAAAGCTGACTTA 419
Db 184 TCAAAACCAATTTGCTACTGATAGTGGCGGATGTCACATATAAATTTGAGAAAGCTGACTTA 243

Qy 420 CTAAGGCTATTCAAGAACAAATTGATCGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 479
Db 244 CTAAGGCTATTCAAGAACAAATTGATCGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 303

Qy 480 GTCAATTGATTTGCAAGCGATCAACCATTTACTGATCGAAACCGCAGGCTCTACTTTGCT 539
Db 304 GTCAATTGATTTGCAAGCGATCAACCATTTACTGATCGAAACCGCAGGCTCTACTTTGCT 363
```

Qy 540 GACAAAGATGGTTCGGTAACTTGGCGAGCCCACTGTCTCAAGAAATTTTGTCTAAGCGGA 599
Db |||||
364 GACAAAGATGGTTCGGTAACTTGGCGAGCCCACTGTCTCAAGAAATTTTGTCTAAGCGGA 423
Qy 600 CATGTGCGGTTAGACCATATAAAGAAAAACCAATACAAACCAAGCGAAATCTGTGTGAT 659
Db |||||
424 CATGTGCGGTTAGACCATATAAAGAAAAACCAATACAAACCAAGCGAAATCTGTGTGAT 483
Qy 660 GTGGGAATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTC 719
Db |||||
484 GTGGGAATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTC 543
Qy 720 AAAGATACCTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTA 779
Db |||||
544 AAAGATACCTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTA 603
Qy 780 CTAGCTCAAGCAAAAGCAATTTTAAACAAAAACCCAGGCTATACGATTTTATGAACGT 839
Db |||||
604 CTAGCTCAAGCAAAAGCAATTTTAAACAAAAACCCAGGCTATACGATTTTATGAACGT 663
Qy 840 GACTCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAA 899
Db |||||
664 GACTCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAA 723
Qy 900 GAGTTTACTTACCGTGTAAAAAATCGGAAACCAAGCTTTATAGGATCAATAAAAAATCTGCT 959
Db |||||
724 GAGTTTACTTACCGTGTAAAAAATCGGAAACCAAGCTTTATAGGATCAATAAAAAATCTGCT 783
Qy 960 CTGAATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAA 1019
Db |||||
784 CTGAATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAA 843
Qy 1020 GGGGAAAGCCGTATGATCCCTTTGATCGAGTCACTTGAACTGTTTACCATCAAAATAC 1079
Db |||||
844 GGGGAAAGCCGTATGATCCCTTTGATCGAGTCACTTGAACTGTTTACCATCAAAATAC 903
Qy 1080 GTTGATGTGATACCAACCAAGTGTGTAAGAGTGACGCTCTTAAACAGCTAGCGAACGT 1139
Db |||||
904 GTTGATGTGATACCAACCAAGTGTGTAAGAGTGACGCTCTTAAACAGCTAGCGAACGT 963
Qy 1140 AACTTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTCTCTACAAACAT 1199
Db |||||
964 AACTTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTCTCTACAAACAT 1023
Qy 1200 CTCGATGCTTTTGGTATTATGGAATATACCTTAACTGGAAAAAGTAGAGGATATACGAT 1259
Db |||||
1024 CTCGATGCTTTTGGTATTATGGAATATACCTTAACTGGAAAAAGTAGAGGATATACGAT 1083
Qy 1260 GACACCAACCGTATCATACCCGTTTATATGGGCAAGCGACCCGAGGAGGATGCTAGC 1319
Db |||||
1084 GACACCAACCGTATCATACCCGTTTATATGGGCAAGCGACCCGAGGAGGATGCTAGC 1143
Qy 1320 TATCATTTAGCCGGTGT 1337
Db |||||
1144 TATCATTTAGCCGTATGAT 1161

RESULT 13
LOCUS E01413 1407 bp DNA linear PAT 29-SEP-1997
DEFINITION DNA sequence of streptokinase.
ACCESSION E01413
VERSION E01413.1 GI:2169669
KEYWORDS JP 1987296881-A/1.
SOURCE Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 1407)
AUTHORS Marii,J.H. and Deebuitsudo,U.S.
TITLE YEAST DNA FRAGMENT AND ITS UTILIZATION

JOURNAL Patent: JP 1987296881-A 1 24-DEC-1987;
COMMENT PHILLIPS PETROLEUM CO
PN JP 1987296881-A/1
PD 24-DEC-1987
PF 02-MAY-1987 JP 1987109620
PR 08-MAY-1986 US 86 860960
PI MARI JIEIN HEIJENSON, DEEBUITSUDO UOMATSUKU SUTOROOWAN PC
C12N15/00, C12N1/16, C12N9/12, (C12N1/16, C12N1/84); CC strandedness:
Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers
FT CDS 8..1255
FT /product='streptokinase'
FT 5'UTR 1..7
FT 3'UTR 1256..1407.
FEATURES
source 1..1407
/organism="Lycopersicon esculentum"
/mol_type="genomic DNA"
/db_xref="taxon:4081"
ORIGIN
Query Match 67.8%; Score 1126.2; DB 6; Length 1407;
Best Local Similarity 98.0%; Pred. No. 1.5e-285; Indels 0; Gaps 0;
Matches 1140; Conservative 0; Mismatches 23;
Qy 175 ATACCATGATTTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAACGCCAAT 234
Db |||||
3 ATTCCATGATTTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAACGCCAAT 62
Qy 235 TGGTTGTTAGCGTTGCTGGTACTGTTGAGGGAGCAATCAAGACATTTAGTCTTAAATTTT 294
Db |||||
63 TAGTTGTTAGCGTTGCTGGTACTGTTGAGGGAGCAATCAAGACATTTAGTCTTAAATTTT 122
Qy 295 TTGAAATCGATCTAAACATCAGACCTGCTCATGAGGAGAAACAGACAGCGCTTAAAGTC 354
Db |||||
123 TTGAAATCGATCTAAACATCAGACCTGCTCATGAGGAGAAACAGACAGCGGTTAAAGTC 182
Qy 355 CAAAAATCAAAACCAATTTGCTACTGATAGTGGCGGATGTACATAAACTTGAGAAAGCTG 414
Db |||||
183 CAAAAATCAAAACCAATTTGCTACTGATAGTGGCGGATGTACATAAACTTGAGAAAGCTG 242
Qy 415 ACTTACTAAAGGCTATTCAAGAACAAATTGATCGCTAAACGCTCCACAGTAAACGACGACTACT 474
Db |||||
243 ACTTACTAAAGGCTATTCAAGAACAAATTGATCGCTAAACGCTCCACAGTAAACGACGACTACT 302
Qy 475 TTGAGGTCATTTGATTTTGCAGCGATGCAACCATTTACTGATCGGAAACGGGAAAGGCTTACT 534
Db |||||
303 TTGAGGTCATTTGATTTTGCAGCGATGCAACCATTTACTGATCGGAAACGGGAAAGGCTTACT 362
Qy 535 TTGCTGACAAAGATGTTTCCGTAACTTGGCGACCAACCTGTCCAAAGAAATTTTGTCTAA 594
Db |||||
363 TTGTCGACAAAGATGTTTCCGTAACTTGGCGACCAACCTGTCCAAAGAAATTTTGTCTAA 422
Qy 595 GCGGACATGTGCGGTTAGACCATATAAAGAAAAACCAATACAAAAACCAAGCGAAATCTG 654
Db |||||
423 GCGGACATGTGCGGTTAGACCATATAAAGAAAAACCAATACAAAAACCAAGCGAAATCTG 482
Qy 655 TTGATGTGGAATATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCG 714
Db |||||
483 TTGATGTGGAATATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCG 542
Qy 715 GTCTCAAGATATCTAAGCTATTGAAAACACTTAGCTATCGGTGACACCATCAATCTCAAG 774
Db |||||
543 GTCTCAAGATATCTAAGCTATTGAAAACACTTAGCTATCGGTGACACCATCAATCTCAAG 602
Qy 775 AATTACTAGCTCAAGCACAAAGCAATTTTAAACAAAAACCCAGGCTTATACGATTTATG 834
Db |||||
603 AATTACTAGCTCAAGCACAAAGCAATTTTAAACAAAAACCCAGGCTTATACGATTTATG 662

```
QY 835 AACGTGACTCTCAATCGTCACTCATGACAAATGACATTTTCGTACGATTTTACCAATGG 894
Db 663 AACGTGACTCTCAATCGTCACTCATGACAAATGACATTTTCGTACGATTTTACCAATGG 722
QY 895 ATCAAGAGTTTACTTACCGTGTAAATAATCGGAAACAAAGCTTATAGGATCAATAAAAT 954
Db 723 ATCAAGAGTTTACTTACCGTGTAAATAATCGGAAACAAAGCTTATAGGATCAATAAAAT 782
QY 955 CTGGTCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTA 1014
Db 783 CTGGTCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTA 842
QY 1015 AAAAAGGGGAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCACCAATCA 1074
Db 843 AAAAAGGGGAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCACCAATCA 902
QY 1075 AATACGTTGATGTCGATACCAAGAAATGCTTAAAGAGTGACGCTCTTAAACGATAGG 1134
Db 903 AATACGTTGATGTCGATACCAAGAAATGCTTAAAGAGTGACGCTCTTAAACGATAGG 962
QY 1135 AACGTAACTTAGACTTTCAGAGATTTATACGATCTCTGATGATAAGCTTAACTTCTTACA 1194
Db 963 AACGTAACTTAGACTTTCAGAGATTTATACGATCTCTGATGATAAGCTTAACTTCTTACA 1022
QY 1195 ACATCTCGATGCTTTTGGTATTATGACTATACCTTAACTTGAAGAAAGTAGAGGATAATC 1254
Db 1023 ACATCTCGATGCTTTTGGTATTATGACTATACCTTAACTTGAAGAAAGTAGAGGATAATC 1082
QY 1255 ACATGACACCAACCGTATATACCGTTTATATGGGCAAGCGACCGAGGAGAGATG 1314
Db 1083 ACATGACACCAACCGTATATACCGTTTATATGGGCAAGCGACCGAGGAGAGATG 1142
QY 1315 CTAGCTATCATTTAGCGGGTGGT 1337
Db 1143 CTAGCTATCATTTAGCTATGAT 1165
```

RESULT 14

```
A20016
LOCUS A20016
DEFINITION A20016 1512 bp DNA linear PAT 14-JUL-1995
          SEQ ID NO: 24; Nucleotide sequence for streptokinase fused to yeast
          alpha-factor.
ACCESSION A20016
VERSION A20016.1 GI:1247850
KEYWORDS synthetic construct
SOURCE synthetic construct
          other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 1512)
          Location/Qualifiers
            1..1512
              /organism="synthetic construct"
              /mol_type="unassigned DNA"
              /db_xref="taxon:32630"
              7..1506
                /codon_start=1
                /transl_table=11
                /product="streptokinase fused to yeast alpha-factor"
                /protein_id="CAA01487.1"
                /db_xref="GI:1247851"
                /translation="MRPPIFTAVLFAASSALAAPVNTTETETAQIPABAVIGYLDL
                EGDVAVLVFSGTNNGLLIFNTTASIAAKEEGVSLDKRIAGPEWLLDRPSWNQSO
                LVSVAGTVGGNODISLAFPRIDLTSPAHGKTEQGLSPKSPKSPATDSGAMPHKLE
                KADLLAKIOELIANVHSDNDVEVIDPASDATIDRNGKVYPADKDGSVLTPTQPVQ
                EFTLSGHRVPRYKPEKPNQAKSVDEVTYVQTPINLPDDDFRPLKDKTKLLKTLAIG
                DITLQELLAQSLINKHPGTIYVERSSIVTHNDNIFRTILPMDQFTYHVKRE
                QAYENKKSGLNEENINTDLISEKTVLKKGEKPPFPORSHLKLFTIKYVDVNTNEL
                LKSEOLLTASERNLDRDLPDRDKALLNNLDAFGIMDYTLTGKVEDNHDNTRI
                TVYMGKRPGENASYHLAYDKORYTEEREVSYLRYTGTPIPDNPDK"
```

CDS

```
QY 895 ATCAAGAGTTTACTTACCGTGTAAATAATCGGAAACAAAGCTTATAGGATCAATAAAAT 954
Db 974 ATCAAGAGTTTACTTACCGTGTAAATAATCGGAAACAAAGCTTATAGGATCAATAAAAT 1033
QY 955 CTGGTCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTA 1014
Db 1034 CTGGTCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTA 1093
QY 1015 AAAAAGGGGAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCACCAATCA 1074
Db 1094 AAAAAGGGGAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCACCAATCA 1153
QY 1075 AATACGTTGATGTCGATACCAAGAAATGCTTAAAGAGTGACGCTCTTAAACGATAGG 1134
Db 1154 AATACGTTGATGTCGATACCAAGAAATGCTTAAAGAGTGACGCTCTTAAACGATAGG 1213
QY 1135 AACGTAACTTAGACTTTCAGAGATTTATACGATCTCTGATGATAAGCTTAACTTCTTACA 1194
Db 1214 AACGTAACTTAGACTTTCAGAGATTTATACGATCTCTGATGATAAGCTTAACTTCTTACA 1273
```

ORIGIN

QY 1195 ACAATCTCGATGCTTTGGTATTATGAGCTATACCTTAACCTGGAAGATGAGAGATAATC 1254
 Db 1274 ACAATCTCGATGCTTTGGTATTATGAGCTATACCTTAACCTGGAAGATGAGAGATAATC 1333
 QY 1255 ACGATGACACCAACCGTATCATACCGTTTATATGCGCAAGCGACCCGAAAGGAGAGATG 1314
 Db 1334 ACGATGACACCAACCGTATCATACCGTTTATATGCGCAAGCGACCCGAAAGGAGAGATG 1393
 QY 1315 CTAGCTATCATATTAGCCGGTGGT 1337
 Db 1394 CTAGCTATCATATTAGCCTATGAT 1416

RESULT 15
 LOCUS I13204 1512 bp DNA linear PAT 26-JUL-1995
 DEFINITION Sequence 27 from patent US 5434073.
 ACCESSION I13204
 VERSION I13204.1 GI:910552
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.
 REFERENCE 1 (bases 1 to 1512)
 AUTHORS Dawson,K., Hunter,M.G. and Czaplowski,L.G.
 TITLE Fibrinolytic and anti-thrombotic cleavable dimers
 JOURNAL Patent: US 5434073-A 27 18-JUL-1995;
 FEATURES
 source Location/Qualifiers
 1..1512
 /organism="unknown"
 /mol_type="unassigned DNA"

ORIGIN
 Query Match 67.8%; Score 1126.2; DB 6; Length 1512;
 Best Local Similarity 98.0%; Pred. No. 1.5e-285;
 Matches 1140; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 175 ATACCATGATTCGTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAACAGCCAAT 234
 Db 254 ATAAAGAAATTCGTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAACAGCCAAT 313
 QY 235 TGGTGTGTTAGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTT 294
 Db 314 TAGTGTGTTAGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTT 373
 QY 295 TTGAATTCGATCTAATCAATCAACGACCTGCTCATGAGGAAAGACAGACAAAGGCTTAAGTC 354
 Db 374 TTGAATTCGATCTAATCAATCAACGACCTGCTCATGAGGAAAGACAGACAAAGGCTTAAGTC 433
 QY 355 CAAATCAAAACGNTTGTCTACTGATAGTGGCGGATGTCAATAACTTGAGAAAGCTG 414
 Db 434 CAAATCAAAACGNTTGTCTACTGATAGTGGCGGATGTCCACATAAACTTGAAGAAAGCTG 493
 QY 415 ACTTACTAAAGGCTATTCAAGAACCAATTCATCCCTAACCGTCCACAGTAAACGACGACTACT 474
 Db 494 ACTTACTAAAGGCTATTCAAGAACCAATTCATCCCTAACCGTCCACAGTAAACGACGACTACT 553
 QY 475 TTGAGTCTCATTTGATTTTCAAGCGATGCAACCAATTAATCTGATCGAAACGGCAAGGCTTACT 534
 Db 554 TTGAGTCTCATTTGATTTTCAAGCGATGCAACCAATTAATCTGATCGAAACGGCAAGGCTTACT 613
 QY 535 TTGCTGACAAAGATGGTTGCGTAAACCTTGGCGACCCCAACCTGTCCAAAGAAATTTTGTCTAA 594
 Db 614 TTGCTGACAAAGATGGTTGCGTAAACCTTGGCGACCCCAACCTGTCCAAAGAAATTTTGTCTAA 673
 QY 595 GCGGACATGTCGCGTTAGACCAATATAAGAAAGAAACCAATACAAACCAAGCGAAATCTG 654
 Db 674 GCGGACATGTCGCGTTAGACCAATATAAGAAAGAAACCAATACAAACCAAGCGAAATCTG 733
 QY 655 TTGATGTGGAATATACTGTACAGTTTACTCCCTTTAAACCCCTGTATGACGATTTTCAGACCAG 714
 Db 734 TTGATGTGGAATATACTGTACAGTTTACTCCCTTTAAACCCCTGTATGACGATTTTCAGACCAG 793

QY 715 GTCTCAAAGATFACCTAAGCTATTGAAAACACTAGCTATCGGTGACACCAATCAATCTCAAG 774
 Db 794 GTCTCAAAGATFACCTAAGCTATTGAAAACACTAGCTATCGGTGACACCAATCAATCTCAAG 853
 QY 775 AATTACTAGCTCAAGCAAAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTATG 834
 Db 854 AATTACTAGCTCAAGCAAAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTATG 913
 QY 835 AACGTGACTCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGG 894
 Db 914 AACGTGACTCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGG 973
 QY 895 ATCAAGAGTTTACTTTACCGTGTAAAAATCGGGAACAAAGCTTTATAGGATCAATAAAAAAT 954
 Db 974 ATCAAGAGTTTACTTTACCATGTCAAAATCGGGAACAAAGCTTTATGAGATCAATAAAAAAT 1033
 QY 955 CTGCTCTGAATGAAGAAATAAAACAACCTGACCTGTATCTCTGAGAAATATTTACGTCCTTA 1014
 Db 1034 CTGCTCTGAATGAAGAAATAAAACAACCTGACCTGTATCTCTGAGAAATATTTACGTCCTTA 1093
 QY 1015 AAAAAGGGGAAAAGCCGCTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCAACATCA 1074
 Db 1094 AAAAAGGGGAAAAGCCGCTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCAACATCA 1153
 QY 1075 AATACGTTGATGTCGATACCAACCAAGAAATGCTTAAAAAGTGAGCAGCTCTTAAACAGCTAGCG 1134
 Db 1154 AATACGTTGATGTCGATACCAACCAAGAAATGCTTAAAAAGTGAGCAGCTCTTAAACAGCTAGCG 1213
 QY 1135 AACGTAACTTAGACTTCAGAGATTTATACGATCTCTCGTGAATAGGCTTAACTACTCTACA 1194
 Db 1214 AACGTAACTTAGACTTCAGAGATTTATACGATCTCTCGTGAATAGGCTTAACTACTCTACA 1273
 QY 1195 ACAATCTCGATGCTTTTGGTATTATGGAATATACCTTAACTGGGAAAGTAGAGGATAATC 1254
 Db 1274 ACAATCTCGATGCTTTTGGTATTATGGAATATACCTTAACTGGGAAAGTAGAGGATAATC 1333
 QY 1255 ACGATGACACCAACCGTATCATACCGTTTATATGCGCAAGCGACCCGAAAGGAGAGATG 1314
 Db 1334 ACGATGACACCAACCGTATCATACCGTTTATATGCGCAAGCGACCCGAAAGGAGAGATG 1393
 QY 1315 CTAGCTATCATTTAGCCGGTGGT 1337
 Db 1394 CTAGCTATCATTTAGCCTATGAT 1416

Search completed: February 1, 2006, 04:17:47
 Job time : 8445.12 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2006, 10:44:08 ; Search time 1004.11 Seconds
(without alignments)
11024.772 Million cell updates/sec

Title: US-09-940-235-10
Perfect score: 1661
Sequence: 1 gcaaccccgccagcctagcc.....gaataagctgtaccatctaa 1661

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0
Searched: 4996997 seqs, 3332346308 residues
Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_21.*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005as:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1661	100.0	1661	3	Aaa37637 Chimeric
2	1282.2	77.2	1377	3	Aaa37622 Streptoki
3	1219.4	73.4	1327	3	Aaa37628 Streptoki
4	1185	71.3	1541	3	Aaa37644 Chimeric
5	1165.8	70.2	2096	3	Aaa37643 Chimeric
6	1150.2	69.2	1782	3	Aaa37642 Chimeric
7	1150	69.2	2385	2	Aax80497 Streptoki
8	1148.8	69.2	2030	2	Aaql1651 FB-PB-SK
9	1148.6	69.2	1242	2	Aax80492 Streptoco
10	1148.6	69.2	1245	3	Aaa37633 S. equisi
11	1148.6	69.2	1254	6	ABA05546 Streptoki
12	1148.6	69.2	8893	6	ABA05547 Maxadilan
13	1147	69.1	1242	2	Aax16632 Streptoco
14	1143.8	68.9	1242	2	Aax16633 Streptoco
15	1136.2	68.4	7057	12	ADM01294 Plasmid p
16	1134.6	68.3	2566	2	Aat77778 Coding se
17	1126.2	67.8	1407	1	Aan70106 DNA encod
18	1126.2	67.8	1512	2	Aaql2158 Streptoki
19	1124.6	67.7	1245	2	AaQ20665 SKC-2 str

20	1124.6	67.7	1335	2	AAQ12156
21	1124.6	67.7	1458	2	AAQ12162
22	1124.6	67.7	1467	2	AAQ12490
23	1124.6	67.7	2589	2	AAQ12160
24	1116.6	67.2	1473	2	AAQ05603
25	1108.6	66.7	2568	1	AAN50493
26	1107	66.6	1323	2	AAT29961
27	1099.4	66.2	1242	5	AAF82144
28	1096.6	66.0	1122	3	AZ299251
29	1096.6	66.0	1158	3	AZ299252
30	1089.2	65.6	1209	3	AZ299249
31	1089.2	65.6	1245	3	AZ299250
32	1080.4	65.0	2253	2	AAQ12161
33	1078.8	64.9	1119	2	AAQ12159
34	995	59.9	1320	6	ABN70192
35	995	59.9	1323	13	ADR83811
36	990.2	59.6	1473	2	AAQ05604
37	976.2	58.8	1068	2	AAX80493
38	971.6	58.5	2208	2	AAX83589
39	899.4	54.1	1245	10	ADF48644
40	762.8	45.9	1262	2	AAQ10230
41	450	27.1	450	2	AAX80494
42	316.8	19.1	4860	3	AAA35009
43	316.8	19.1	6988	9	ACD06169
44	316.8	19.1	7020	13	ADR90518
45	316.8	19.1	7049	12	ADN95947

ALIGNMENTS

RESULT 1

AAA37637	AAA37637 standard; DNA; 1661 BP.
XX	AC AAA37637;
XX	15-SEP-2003 (revised)
DT	DT 13-OCT-2000 (first entry)
XX	Chimeric SK-FBD coding sequence.
DE	Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
KW	plasminogen; human; fibrinectin; thrombolytic therapy;
KW	cardiovascular disorder; ss.
XX	Streptococcus dysgalactiae subsp. equisimilis.
OS	Homo sapiens.
OS	Chimeric.
XX	EPI024192-A2.
XX	02-AUG-2000.
XX	23-DEC-1999; 99EP-00310541.
XX	24-DEC-1998; 98IN-DE003825.
XX	(COUL) CSIR COUNCIL SCI IND RES.
XX	Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
PI	Yadav M;
XX	WPI; 2000-516032/47.
DR	Hybrid streptokinase-fibrin binding domain polypeptides useful for
XX	thrombolytic therapy comprises a streptokinase fused with fibrin binding
PT	domains of human fibrinectin.
XX	Example 5; Fig 19b; 58pp; English.
XX	This sequence represents a chimeric streptokinase-fibrin binding domain
CC	(SK-FBD) protein coding sequence. The invention relates to a hybrid

CC plasminogen activator (PA) comprises a polypeptide fusion between
 CC streptokinase (SK), which are capable of plasminogen (PG) activation, and
 CC fibrin binding regions of human fibronectin, which are from fibrin
 CC binding domains (PBD) 4 and 5 or 1 and 2. The hybrid PA possesses the
 CC ability to bind with fibrin independently and also characteristically
 CC retains a PG activation ability which becomes evident only after a
 CC pronounced duration, or lag, after exposure of the PA to a suitable
 CC animal or human PG. The hybrid streptokinase-fibrin binding domain
 CC polypeptides are useful in thrombolytic therapy for various kinds of
 CC cardiovascular disorders. The hybrids have enhanced fibrin selectivity as
 CC well as kinetics of plasminogen activation that are distinct from that of
 CC natural streptokinase in being characterised by a temporary delay, or lag
 CC of several minutes in the natural rate of the catalytic conversion of
 CC plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins
 CC can bind tightly with fibrin in blood clots soon after introduction into
 CC the vascular system without significantly activating the circulating
 CC blood plasminogen to plasmin, thus aiding in the localisation of the
 CC plasminogen activation process to the site of pathological thrombus. This
 CC overcomes systemic plasminogen activation encountered during clinical use
 CC of streptokinase. (Updated on 15-SEP-2003 to standardise OS field)

XX Sequence 1661 BP; 529 A; 370 C; 357 G; 405 T; 0 U; 0 Other;

Query Match 100.0%; Score 1661; DB 3; Length 1661;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1661; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCAACCCCGCGAGCTAGCGGGTCTCAACGACAGGACGATCATGCGACCCCGTGG 60
 DB 1 GCAACCCCGCGAGCTAGCGGGTCTCAACGACAGGACGATCATGCGACCCCGTGG 60
 QY 61 CCAGGACCAACGCTGCGCGAGATCGATCCCGGAAATTAATACGACTCACTATAGG 120
 DB 61 CCAGGACCAACGCTGCGCGAGATCGATCCCGGAAATTAATACGACTCACTATAGG 120
 QY 121 AGACACCAACGCTTCCCTCTAGAAATAATTTTGTGTTAACTTTAAGAGGAGATATACCA 180
 DB 121 AGACACCAACGCTTCCCTCTAGAAATAATTTTGTGTTAACTTTAAGAGGAGATATACCA 180
 QY 181 TGAATGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAGCAATTTGGTTG 240
 DB 181 TGAATGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAGCAATTTGGTTG 240
 QY 241 TTAGCGTTGCTGGTACTGTGAGGGACGATCAAGACATAGTCTTAATTTTTTGAAA 300
 DB 241 TTAGCGTTGCTGGTACTGTGAGGGACGATCAAGACATAGTCTTAATTTTTTGAAA 300
 QY 301 TCGATCTTAACATCAGCACTGCTCATGGAGAAAGACAGACGAGGCTTAAGTCCAAAT 360
 DB 301 TCGATCTTAACATCAGCACTGCTCATGGAGAAAGACAGACGAGGCTTAAGTCCAAAT 360
 QY 361 CAAAACCAATTTGCTACTGATAGTGGCGCGATGTACATAACTTTGAGAAAGCTGACTTAC 420
 DB 361 CAAAACCAATTTGCTACTGATAGTGGCGCGATGTACATAACTTTGAGAAAGCTGACTTAC 420
 QY 421 TAAAGGCTATTCAAGAACCAATTTGATCGCTAACGTCACAGTAAACGACGACTACTTTGAGG 480
 DB 421 TAAAGGCTATTCAAGAACCAATTTGATCGCTAACGTCACAGTAAACGACGACTACTTTGAGG 480
 QY 481 TCATTGATTTTGAAGCGGATCAACCATTAATCTGATCGAAACGGCAAGGCTACTTTGCTG 540
 DB 481 TCATTGATTTTGAAGCGGATCAACCATTAATCTGATCGAAACGGCAAGGCTACTTTGCTG 540
 QY 541 ACAAGAGTGGTTCGGTAACTTGGCCGACCAACCTGTCCAAAGAAATTTTGTCTAAGCGGAC 600
 DB 541 ACAAGAGTGGTTCGGTAACTTGGCCGACCAACCTGTCCAAAGAAATTTTGTCTAAGCGGAC 600
 QY 601 ATGTGGCGTTAGACCATATAAAGAAACCAATATAAAGCAAGCAATCTGTTGATG 660
 DB 601 ATGTGGCGTTAGACCATATAAAGAAACCAATATAAAGCAAGCAATCTGTTGATG 660
 QY 661 TCGAATATATCTGACAGTTTACTCCCTTAACACCTGATGACGATTTTCAGACCAAGTCTCA 720
 DB 661 TCGAATATATCTGACAGTTTACTCCCTTAACACCTGATGACGATTTTCAGACCAAGTCTCA 720

DB 661 TCGAATATATCTGACAGTTTACTCCCTTAACACCTGATGACGATTTTCAGACCAAGTCTCA 720
 QY 721 AAGATCTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATTAC 780
 DB 721 AAGATCTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATTAC 780
 QY 781 TAGCTCAAGCACAAAGCAATTTAAACAAAACACCCAGGCTATACGATTTTATGAACGTG 840
 DB 781 TAGCTCAAGCACAAAGCAATTTAAACAAAACACCCAGGCTATACGATTTTATGAACGTG 840
 QY 841 ACTCTCAATCGTCACTCATGACAAATGCAATTTTCGTCGACGATTTTACCAATGATCAAG 900
 DB 841 ACTCTCAATCGTCACTCATGACAAATGCAATTTTCGTCGACGATTTTACCAATGATCAAG 900
 QY 901 AGTTTACTTACCGTGTAAAAATCGGAAACGACTTATAGGATCAATAAAAAATCTGGTC 960
 DB 901 AGTTTACTTACCGTGTAAAAATCGGAAACGACTTATAGGATCAATAAAAAATCTGGTC 960
 QY 961 TGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTTACGTCCTTAAAAAAG 1020
 DB 961 TGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTTACGTCCTTAAAAAAG 1020
 QY 1021 GGGAAAAGCCGTATGATCCCTTTTGATCGCAGTCACTTTGAAACTGTTCAACATCAAAATACG 1080
 DB 1021 GGGAAAAGCCGTATGATCCCTTTTGATCGCAGTCACTTTGAAACTGTTCAACATCAAAATACG 1080
 QY 1081 TTGATGTCGATACCAACGAATTTGCTAAAAAGTGAAGAGTCTTTAAACAGCTAGCGAACGTA 1140
 DB 1081 TTGATGTCGATACCAACGAATTTGCTAAAAAGTGAAGAGTCTTTAAACAGCTAGCGAACGTA 1140
 QY 1141 ACTTAGACTTTTCTGAGATTTATACGATCTCTGATTAAGGCTTAACTCTCTCAACAAATC 1200
 DB 1141 ACTTAGACTTTTCTGAGATTTATACGATCTCTGATTAAGGCTTAACTCTCTCAACAAATC 1200
 QY 1201 TCGATGCTTTTGGTATTATGACATATACCTTTAACTGGAAGAGTAGAGGATTAATCACGATG 1260
 DB 1201 TCGATGCTTTTGGTATTATGACATATACCTTTAACTGGAAGAGTAGAGGATTAATCACGATG 1260
 QY 1261 ACACCAACCGTATCATTAACCGTTTATATGGAAGAGGACCCGAAAGGAGAGATGCTAGCT 1320
 DB 1261 ACACCAACCGTATCATTAACCGTTTATATGGAAGAGGACCCGAAAGGAGAGATGCTAGCT 1320
 QY 1321 ATCAATTTAGCCGGTGGTGGTCAAGCGAGCAAAATGTTTCAGCCCCAGTCCCGGTTGGCTG 1380
 DB 1321 ATCAATTTAGCCGGTGGTGGTCAAGCGAGCAAAATGTTTCAGCCCCAGTCCCGGTTGGCTG 1380
 QY 1381 TCAGTCAAAAGCAAGCCCGGTTGTTATGACAAATGGAAGAACATATCAGATAAATCAACAGT 1440
 DB 1381 TCAGTCAAAAGCAAGCCCGGTTGTTATGACAAATGGAAGAACATATCAGATAAATCAACAGT 1440
 QY 1441 GGGAGCGGACCTACTAGGTAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1500
 DB 1441 GGGAGCGGACCTACTAGGTAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1500
 QY 1501 TTAACCTCGGAAAGTAAACCTGAAAGTGAAGAGACTTCTGTTTGAAGTCACTGCGGAC 1560
 DB 1501 TTAACCTCGGAAAGTAAACCTGAAAGTGAAGAGACTTCTGTTTGAAGTCACTGCGGAC 1560
 QY 1561 CTTACCGAGTGGGTGACACTTATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACCT 1620
 DB 1561 CTTACCGAGTGGGTGACACTTATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACCT 1620
 QY 1621 GCATCGGGGCTGGCGAGGAGAAATAAGCTGTATCATCTAA 1661
 DB 1621 GCATCGGGGCTGGCGAGGAGAAATAAGCTGTATCATCTAA 1661

RESULT 2
 AAA37622
 ID AAA37622 standard; DNA; 1377 BP.
 XX
 AC AAA37622;
 XX

DT 15-SEP-2003 (revised)
 DT 13-OCT-2000 (first entry)
 XX Streptokinase-NTRN gene.
 DE
 DE Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
 KW plasminogen; human; fibrinectin; thrombolytic therapy;
 KW cardiovascular disorder; ss.
 XX
 OS Streptococcus dysgalactiae subsp. equisimilis.
 XX
 XX EP1024192-A2.
 PN
 XX
 PD 02-AUG-2000.
 XX
 XX 23-DEC-1999; 99EP-00310541.
 XX
 XX 24-DEC-1998; 98IN-DE003825.
 PR
 XX (COUL) CSIR COUNCIL SCI IND RES.
 PA
 XX Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
 PI Yadav M;
 PI
 XX WPI; 2000-516032/47.
 DR
 XX Hybrid streptokinase-fibrin binding domain polypeptides useful for
 PT thrombolytic therapy comprises a streptokinase fused with fibrin binding
 PT domains of human fibrinectin.
 XX
 XX Example 1; Fig 11; 58pp; English.
 PS
 XX This sequence represents a streptokinase-NTRN (SK-NTRN) gene (where NTRN
 CC stands for N-terminally repaired with native sequence). The invention
 CC relates to a hybrid plasminogen activator (PA) comprises a polypeptide
 CC fusion between streptokinase (SK), which are capable of plasminogen (PG)
 CC activation, and fibrin binding regions of human fibrinectin, which are
 CC from fibrin binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA
 CC possesses the ability to bind with fibrin independently and also
 CC characteristically retains a PG activation ability which becomes evident
 CC only after a pronounced duration, or lag, after exposure of the PA to a
 CC suitable animal or human PG. The hybrid streptokinase-fibrin binding
 CC domain polypeptides are useful in thrombolytic therapy for various kinds
 CC of cardiovascular disorders. The hybrids have enhanced fibrin selectivity
 CC as well as kinetics of plasminogen activation that are distinct from that
 CC of natural streptokinase in being characterised by a temporary delay, or
 CC lag of several minutes in the natural rate of the catalytic conversion of
 CC plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins
 CC can bind tightly with fibrin in blood clots soon after introduction into
 CC the vascular system without significantly activating the circulating
 CC blood plasminogen to plasmin, thus aiding in the localisation of the
 CC plasminogen activation process to the site of pathological thrombus. This
 CC overcomes systemic plasminogen activation encountered during clinical use
 CC of streptokinase. (Updated on 15-SEP-2003 to standardise OS field)
 XX
 SQ Sequence 1377 BP; 465 A; 302 C; 265 G; 345 T; 0 U; 0 Other;
 Query Match 77.2%; Score 1282.2; DB 3; Length 1377;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1284; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 51 GCACCGGTGGCCAGACCAACGCTCCCGAGATCTCGATCCCGGAAATTAATACGACT 110
 DB 1 GCACCGGTGGCCAGGACCAACGCTCCCGAGATCTCGATCCCGGAAATTAATACGACT 60
 QY 111 CACTATAGGAGACCAACACGGTTCCCTCTAGATAATTTTGTCTTAACCTTTAAGAGG 170
 DB 61 CACTATAGGAGACCAACACGGTTTCCCTCTAGATAATTTTGTCTTAACCTTTAAGAGG 120
 QY 171 AGATATACCATGATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACCAACAGC 230
 DB 121 AGATATACCATGATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACCAACAGC 180

QY 231 CAATTGGTTGTTAGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAA 290
 DB |||||
 QY 181 CAATTGGTTGTTAGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAA 240
 DB |||||
 QY 291 TTTTGTGAATCGATCTAAACATCAGCAGCTGCTCATGGAGGAAACAGACAGCAAGGCTTAA 350
 DB |||||
 QY 241 TTTTGTGAATCGATCTAAACATCAGCAGCTGCTCATGGAGGAAACAGACAGCAAGGCTTAA 300
 DB |||||
 QY 351 AGTCCAAAATCAAAAACCAATTTGCTACTGATAGTGGCGGATGTCATCATATAAATCTTCAGAAA 410
 DB |||||
 QY 301 AGTCCAAAATCAAAAACCAATTTGCTACTGATAGTGGCGGATGTCATCATATAAATCTTCAGAAA 360
 DB |||||
 QY 411 GCTGACTTTACTAAAGGCTATTCAAGAACAAATTTGATCGCTAAAGCTGACAGTAAACGACGAC 470
 DB |||||
 QY 361 GCTGACTTTACTAAAGGCTATTCAAGAACAAATTTGATCGCTAAAGCTGACAGTAAACGACGAC 420
 DB |||||
 QY 471 TACTTTGAGGTCATTGATTTTGGCAAGCGATGCAACCATTTACTGATCGTAAACCGCAAGGTC 530
 DB |||||
 QY 421 TACTTTGAGGTCATTGATTTTGGCAAGCGATGCAACCATTTACTGATCGTAAACCGCAAGGTC 480
 DB |||||
 QY 531 TACTTTGCTGCAAAAGATGGTTCCGTTAACTTCCGTAACCTTCCGACCCCAACCTGTCCAGAAATTTTGG 590
 DB |||||
 QY 481 TACTTTGCTGCAAAAGATGGTTCCGTTAACTTCCGTAACCTTCCGACCCCAACCTGTCCAGAAATTTTGG 540
 DB |||||
 QY 591 CTAAGCGGACATGTGCGGTTAGACCATATATAAGAAAAACCAATACAAAAACCAAGCGAAA 650
 DB |||||
 QY 541 CTAAGCGGACATGTGCGGTTAGACCATATATAAGAAAAACCAATACAAAAACCAAGCGAAA 600
 DB |||||
 QY 651 TCTGTTGATGTGGAATATATCTGATACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCA 710
 DB |||||
 QY 601 TCTGTTGATGTGGAATATATCTGATACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCA 660
 DB |||||
 QY 711 CCAGGTCCTCAAGATFACCTAACTATTGAAAAACACTAGCTATCGGTGACACCATCATCT 770
 DB |||||
 QY 661 CCAGGTCCTCAAGATFACCTAACTATTGAAAAACACTAGCTATCGGTGACACCATCATCT 720
 DB |||||
 QY 771 CAAGAAATTTACTAGCTCAAGCAACAAGCAATTTTAAACAAAAACCAACCCAGGCTTATACGATT 830
 DB |||||
 QY 721 CAAGAAATTTACTAGCTCAAGCAACAAGCAATTTTAAACAAAAACCAACCCAGGCTTATACGATT 780
 DB |||||
 QY 831 TATGAACGTGACTCTCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCA 890
 DB |||||
 QY 781 TATGAACGTGACTCTCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCA 840
 DB |||||
 QY 891 ATGATCAAGATTTACTTACCGTGTAAATAATCGGAAACAAAGCTTATAGGATCAATAAA 950
 DB |||||
 QY 841 ATGATCAAGATTTACTTACCGTGTAAATAATCGGAAACAAAGCTTATAGGATCAATAAA 900
 DB |||||
 QY 951 AAATCTGCTCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTTACGTC 1010
 DB |||||
 QY 901 AAATCTGCTCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTTACGTC 960
 DB |||||
 QY 1011 CTTAAAAAAGGGGAAAAAGCCGTATGATCCCTTTCATCGCAGTCACTTCAAACTGTTTACC 1070
 DB |||||
 QY 961 CTTAAAAAAGGGGAAAAAGCCGTATGATCCCTTTCATCGCAGTCACTTCAAACTGTTTACC 1020
 DB |||||
 QY 1071 ATCAAAATACGTTGATGTCGATACCAACGAATTTGCTTAAATAAGTGAGCAGCTCTTAAACAGCT 1130
 DB |||||
 QY 1021 ATCAAAATACGTTGATGTCGATACCAACGAATTTGCTTAAATAAGTGAGCAGCTCTTAAACAGCT 1080
 DB |||||
 QY 1131 AGCGAAGCTTAATCTTAGACTTTCAGAGATTTATACGATCTCTGATTAAGGCTTAAACTACTC 1190
 DB |||||
 QY 1081 AGCGAAGCTTAATCTTAGACTTTCAGAGATTTATACGATCTCTGATTAAGGCTTAAACTACTC 1140
 DB |||||
 QY 1191 TACAACAATCTCGATGCTTTTGGTATTATGACCTATACCTTAACTTGAAGAAATGAGGAT 1250
 DB |||||
 QY 1141 TACAACAATCTCGATGCTTTTGGTATTATGACCTATACCTTAACTTGAAGAAATGAGGAT 1200
 DB |||||
 QY 1251 AATCAGATGACCAACCGCTATCATACCGTTTATATGGCAACGACCGCAAGGAGAG 1310
 DB |||||
 QY 1201 AATCAGATGACCAACCGCTATCATACCGTTTATATGGCAACGACCGCAAGGAGAG 1260
 DB |||||
 QY 1311 AATGCTAGCTATCATTTTAGCCGGTGGT 1337

1261 AATGCTAGCTATCATTTAGCCTATGAT 1287

|||||

RESULT 3

AAA37628

ID AAA37628 standard; DNA; 1327 BP.

XX AC AAA37628;

XX 15-SEP-2003 (revised)

DT 13-OCT-2000 (first entry)

XX KW Streptokinase-NTR gene.

DE Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;

KW plasminogen; human; fibrinectin; thrombolytic therapy;

KW cardiovascular disorder; ss.

XX OS Streptococcus dysgalactiae subsp. equisimilis.

PN EP1024192-A2.

XX 02-AUG-2000.

XX 23-DEC-1999; 99EP-00310541.

XX 24-DEC-1998; 98IN-DE003825.

XX (COUL) CSIR COUNCIL SCI IND RES.

XX Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;

PI Yadav M;

XX WPI; 2000-516032/47.

DR Hybrid streptokinase-fibrin binding domain polypeptides useful for

PT thrombolytic therapy comprises a streptokinase fused with fibrin binding

PT domains of human fibrinectin.

XX Example 1; Fig 14; 58pp; English.

PS This sequence represents a streptokinase-NTR (SK-NTR) gene (where NTR

CC stands for N-terminally repaired with native sequence). The invention

CC relates to a hybrid plasminogen activator (PA) comprises a polypeptide

CC fusion between streptokinase (SK), which are capable of plasminogen (PG)

CC activation, and fibrin binding regions of human fibrinectin, which are

CC from fibrin binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA

CC possesses the ability to bind with fibrin independently and also

CC characteristically retains a PG activation ability which becomes evident

CC only after a pronounced duration, or lag, after exposure of the PA to a

CC suitable animal or human PG. The hybrid streptokinase-fibrin binding

CC domain polypeptides are useful in thrombolytic therapy for various kinds

CC of cardiovascular disorders. The hybrids have enhanced fibrin selectivity

CC as well as kinetics of plasminogen activation that are distinct from that

CC of natural streptokinase in being characterised by a temporary delay, or

CC lag of several minutes in the natural rate of the catalytic conversion of

CC plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins

CC can bind tightly with fibrin in blood clots soon after introduction into

CC the vascular system without significantly activating the circulating

CC blood plasminogen to plasmin, thus aiding in the localisation of the

CC plasminogen activation process to the site of pathological thrombus. This

CC overcomes systemic plasminogen activation encountered during clinical use

CC of streptokinase. (Updated on 15-SEP-2003 to standardise OS field)

XX Sequence 1327 BP; 456 A; 279 C; 250 G; 342 T; 0 U; 0 Other;

SQ

Query Match 73.4%; Score 1219.4; DB 3; Length 1327;

Best Local Similarity 99.1%; Pred. No. 0;

Matches 1226; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

101 TAATAGGACTCACTAGGAGACCAACGGTTTCCTCTAGAAATAATTTGTTTAAAC 160

|||||

Db 1 TAATA CGACTCACTATATAGGAGACCAACACGGTTTCCCTCTAGAAATAATTTTGTTTAAC 60

QY 161 TTTAAGAGGAGATATACCATGATTTGCTGGAGCTTCAGTGGCTGTAGACCGTCCATCTCTG 220

Db 61 TTTAAGAGGAGATATACCATGATAGCTGGTCTCTGAATGGCTACTAGATCGTCTCTCTG 120

QY 221 CAACAAACAGCCAAATTTGTTGTTAGCGTTGCTGGTACTGTTGAGGGGAGCAATCAAGACAT 280

Db 121 AAATAACAGCCAAATTTGTTGTTAGCGTTGCTGGTACTGTTGAGGGGAGCAATCAAGACAT 180

QY 281 TAGTCTTAAATTTTTTGAATTCGATCTAAACATACGACCTGCTCATGAGGAGAAAGACAGA 340

Db 181 TAGTCTTAAATTTTTTGAATTCGATCTAAACATACGACCTGCTCATGAGGAGAAAGACAGA 240

QY 341 GCAGGGCTTAAGTCCAAATCAAAACCAATTTGCTACTGATAGTGGCGGATGTCATATAA 400

Db 241 GCAAGGGCTTAAGTCCAAATCAAAACCAATTTGCTACTGATAGTGGCGGATGTCATATAA 300

QY 401 ACTTGAGAAAGCTGACTTACTAAAGGCTATTCAAGAACAAATTTGATCGCTTAACGCTCCACAG 460

Db 301 ACTTGAGAAAGCTGACTTACTAAAGGCTATTCAAGAACAAATTTGATCGCTTAACGCTCCACAG 360

QY 461 TAACGACGACTACTTTGAGGTCATTGATTTTTCAGAGCGATGCAACCAATTTACTGATCGAAA 520

Db 361 TAACGACGACTACTTTGAGGTCATTGATTTTTCAGAGCGATGCAACCAATTTACTGATCGAAA 420

QY 521 CGGCAAGGTCATCTTTGCTGACAAAGATGGTTGGTAACTTGCCTGCGGACCAACCTGTCCA 580

Db 421 CGGCAAGGTCATCTTTGCTGACAAAGATGGTTGGTAACTTGCCTGCGGACCAACCTGTCCA 480

QY 581 AGAATTTTGTAAAGCGGACATGTCGGCTTAGACCATATAAAGAAACCAATACAAAA 640

Db 481 AGAATTTTGTAAAGCGGACATGTCGGCTTAGACCATATAAAGAAACCAATACAAAA 540

QY 641 CCAAGCGAAATCTGTTGATGTGGAATATATCTGACAGTTTACTCCCTTTAAACCCCTGATGA 700

Db 541 CCAAGCGAAATCTGTTGATGTGGAATATATCTGACAGTTTACTCCCTTTAAACCCCTGATGA 600

QY 701 CGATTTTCAGACAGGTCCTCAAGATATACTAAGCTATTGAAACACACTAGCTATCGGTGACAC 760

Db 601 CGATTTTCAGACAGGTCCTCAAGATATACTAAGCTATTGAAACACACTAGCTATCGGTGACAC 660

QY 761 CATCACATCTCAAGAAATTTACTAGCTCAAGCACAAAGCAATTTTAAACAAAAACCCACCG 820

Db 661 CATCACATCTCAAGAAATTTACTAGCTCAAGCACAAAGCAATTTTAAACAAAAACCCACCG 720

QY 821 CTATACGATTTTATGAACCTGACTCTCTCAATCGTCACTCATGCAATGACATTTTCGGTAC 880

Db 721 CTATACGATTTTATGAACCTGACTCTCTCAATCGTCACTCATGCAATGACATTTTCGGTAC 780

QY 881 GATTTTACCAATGAGATCAAGAGTTTACTTACCGTGTAAATAATCGGGAAACAAAGCTTTATAG 940

Db 781 GATTTTACCAATGAGATCAAGAGTTTACTTACCGTGTAAATAATCGGGAAACAAAGCTTTATAG 840

QY 941 GATCAATTAATAAATCTGCTCTGAATGAGAAATTAACCAACTGACCTGATCTCTGAGAA 1000

Db 841 GATCAATTAATAAATCTGCTCTGAATGAGAAATTAACCAACTGACCTGATCTCTGAGAA 900

QY 1001 ATATTACGCTCTTAAATAAGGGGAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTTGAA 1060

Db 901 ATATTACGCTCTTAAATAAGGGGAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTTGAA 960

QY 1061 ACTGTTTCAACCAATACGTTGATGTCGATCCCAACGAAATGCTTAAATAAGGTGAGCAGCT 1120

Db 961 ACTGTTTCAACCAATACGTTGATGTCGATCCCAACGAAATGCTTAAATAAGGTGAGCAGCT 1020

QY 1121 CTTTAACAGCTAGCGAAACGTAATCTAGACTTCAGAGATTTATACGATCTCTCGGTGATAGGC 1180

Db 1021 CTTTAACAGCTAGCGAAACGTAATCTAGACTTCAGAGATTTATACGATCTCTCGGTGATAGGC 1080

QY 1181 TAAACTACTCTTACCAACCAATCTCGATGCTTTTGGTATTATGGAATACCTTAACCTGGA 1240

Db 1081 TAAACTACTCTTACCAACCAATCTCGATGCTTTTGGTATTATGGAATACCTTAACCTGGA 1140

QY 1241 ACTAGAGGATATACAGATGACACCAACCGTATCATAAACCGTTATATGGCGAGCGACC 1300
 Db |||||
 1141 ACTAGAGGATATACAGATGACACCAACCGTATCATAAACCGTTATATGGCGAGCGACC 1200
 QY 1301 CGAAGGAGAGATGCTAGCTATCATATTAGCGGTGGT 1337
 Db |||||
 1201 CGAAGGAGAGATGCTAGCTATCATATTAGCGCTATGAT 1237

RESULT 4
 AAA37644
 ID AAA37644 standard; DNA; 1541 BP.
 XX
 AC AAA37644;
 XX
 DT 15-SEP-2003 (revised)
 DT 13-OCT-2000 (first entry)
 XX
 DE Chimeric SK-FBD coding sequence.
 XX
 KW Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
 KW plasminogen; human; fibrinectin; thrombolytic therapy;
 KW cardiovascular disorder; ss.
 XX
 OS Streptococcus dysgalactiae subsp. equisimilis.
 OS Homo sapiens.
 OS Chimeric.
 XX
 PN EP1024192-A2.
 XX
 PD 02-AUG-2000.
 XX
 PF 23-DEC-1999; 99EP-00310541.
 XX
 PR 24-DEC-1998; 98IN-DE003825.
 XX
 PA (COUL) CSIR COUNCIL SCI IND RES.
 XX
 PI Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
 PI Yadav M;
 XX
 DR WPI; 2000-516032/47.
 XX
 XX Hybrid streptokinase-fibrin binding domain polypeptides useful for
 PT thrombolytic therapy comprises a streptokinase fused with fibrin binding
 PT domains of human fibrinectin.
 PS
 PS Disclosure; Fig 17b; 58pp; English.
 XX
 XX This sequence represents a chimeric streptokinase-fibrin binding domain
 CC (SK-FBD) protein coding sequence. The invention relates to a hybrid
 CC plasminogen activator (PA) comprises a polypeptide fusion between
 CC streptokinase (SK), which are capable of plasminogen (PG) activation, and
 CC fibrin binding regions of human fibrinectin, which are from fibrin
 CC binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA possesses the
 CC ability to bind with fibrin independently and also characteristically
 CC retains a PG activation ability which becomes evident only after a
 CC pronounced duration, or lag, after exposure of the PA to a suitable
 CC animal or human PG. The hybrid streptokinase-fibrin binding domain
 CC polypeptides are useful in thrombolytic therapy for various kinds of
 CC cardiovascular disorders. The hybrids have enhanced fibrin selectivity as
 CC well as kinetics of plasminogen activation that are distinct from that of
 CC natural streptokinase in being characterised by a temporary delay, or lag
 CC of several minutes in the natural rate of the catalytic conversion of
 CC plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins
 CC can bind tightly with fibrin in blood clots soon after introduction into
 CC the vascular system without significantly activating the circulating
 CC blood plasminogen to plasmin, thus aiding in the localisation of the
 CC plasminogen activation process to the site of pathological thrombus. This
 CC overcomes systemic plasminogen activation encountered during clinical use
 CC of streptokinase. (Updated on 15-SEP-2003 to standardise OS field)

SQ Sequence 1541 BP; 497 A; 328 C; 335 G; 381 T; 0 U; 0 Other;
 Query Match 71.3%; Score 1185; DB 3; Length 1541;
 Best Local Similarity 98.8%; Pred. No. 0;
 Matches 1194; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
 QY 151 TTTGTTTAACTTTAAGAGGAGATATACCATGATAGTGGTCTCTGAATGGCTACTAGATC 210
 Db 1 TTTGTTTAACTTTAAGAGGAGATATACCATGATAGTGGTCTCTGAATGGCTACTAGATC 60
 QY 211 GTCCATCTGTCAACACAGGCCAAATTTGGTTGTAGCGTTGCTGGTATCTGTTAGGGGACGA 270
 Db |||||
 61 GTCTTCTGTAAATAAACAGCCAAATTTGGTTGTAGCGTTGCTGGTACTGTTGAGGGGACGA 120
 QY 271 ATCAAGCATTAGTCTTAAATTTTTTGAATTCGATCTAACATCAGCATCTGCTCATGGAG 330
 Db |||||
 121 ATCAAGCATTAGTCTTAAATTTTTTGAATTCGATCTAACATCAGCATCTGCTCATGGAG 180
 QY 331 GAAAGACAGAGCAAGGCTTAAAGTCCAAATCAAAACCAATTTGCTACTGATAGTGGCGGA 390
 Db |||||
 181 GAAAGACAGAGCAAGGCTTAAAGTCCAAATCAAAACCAATTTGCTACTGATAGTGGCGGA 240
 QY 391 TGTCACTATAAATTGAGAGAACTGACTTTAAAGGCTATTCAAGAACTTTGATGCTGCTA 450
 Db |||||
 241 TGTCACTATAAATTGAGAGAACTGACTTTAAAGGCTATTCAAGAACTTTGATGCTGCTA 300
 QY 451 ACGTCCACAGTAACGAGCACTACTTTGAGGTCATTGATTTTGCACGATGCAACCATTA 510
 Db |||||
 301 ACGTCCACAGTAACGAGCACTACTTTGAGGTCATTGATTTTGCACGATGCAACCATTA 360
 QY 511 CTGATCGAAACCGCAAGGCTTACTTTGCTGCAAAAGATGGTTTCGGTAACTTTGCCGCC 570
 Db |||||
 361 CTGATCGAAACCGCAAGGCTTACTTTGCTGCAAAAGATGGTTTCGGTAACTTTGCCGCC 420
 QY 571 AACCTGTCCAAGAAATTTTGTAAAGCGGACATGTGCGGTTAGACCATATAAAGAAAC 630
 Db |||||
 421 AACCTGTCCAAGAAATTTTGTAAAGCGGACATGTGCGGTTAGACCATATAAAGAAAC 480
 QY 631 CAATACAAAACCAAGCGAAATCTGTTGATGTGGAAATATCTGTTACAGTTTACTCCTTAA 690
 Db |||||
 481 CAATACAAAACCAAGCGAAATCTGTTGATGTGGAAATATCTGTTACAGTTTACTCCTTAA 540
 QY 691 ACCCTGATGACGATTTTCAGACCGAGTCTCAAAGATATCTAAAGTATTTGAAACACATGCTA 750
 Db |||||
 541 ACCCTGATGACGATTTTCAGACCGAGTCTCAAAGATATCTAAAGTATTTGAAACACATGCTA 600
 QY 751 TCGGTGACACCATCACATCTCAAGAAATTTACTGATGTGGAATATCTGTTACAGTTTACTCCTTAA 810
 Db |||||
 601 TCGGTGACACCATCACATCTCAAGAAATTTACTGATGTGGAATATCTGTTACAGTTTACTCCTTAA 660
 QY 811 ACCACCGAGGTATACGATTTTAAAGCGTGAATCTCTCAATCGTCACTCATGACCAATGACA 870
 Db |||||
 661 ACCACCGAGGTATACGATTTTAAAGCGTGAATCTCTCAATCGTCACTCATGACCAATGACA 720
 QY 871 TTTTCCGTGAGATTTTACCAATGATCAAGAGTTTACTTACCGTGTGTTAAATCGGAC 930
 Db |||||
 721 TTTTCCGTGAGATTTTACCAATGATCAAGAGTTTACTTACCGTGTGTTAAATCGGAC 780
 QY 931 AAGCTTATAGGATCAATATAAATAATCTGCTCTGAATGAAGAAATATAAACAACGATGCTA 990
 Db |||||
 781 AAGCTTATAGGATCAATATAAATAATCTGCTCTGAATGAAGAAATATAAACAACGATGCTA 840
 QY 991 TCTCTGAGAAATATTACGTCCTTTAAAGGGGAAAGCGGTATGATCCCTTTCGATCGCA 1050
 Db |||||
 841 TCTCTGAGAAATATTACGTCCTTTAAAGGGGAAAGCGGTATGATCCCTTTCGATCGCA 900
 QY 1051 GTCACTTGAACCTGTTCCACCATCAATGATGTGATGCTGATACCAACGATTTGCTAATA 1110
 Db |||||
 901 GTCACTTGAACCTGTTCCACCATCAATGATGTGATGCTGATGCTGATGCTGATGCTGAT 960
 QY 1111 GTGAGGAGCTCTTAAACAGCTAGCGAACTGATGATGATGATGATGATGATGATGATGAT 1170
 Db |||||
 961 GTGAGGAGCTCTTAAACAGCTAGCGAACTGATGATGATGATGATGATGATGATGATGAT 1020

1171 GTGATAAGGCTAACTACTCTACAACTCTCGATGCTTTTGGTATTATGGACTATACCT 1230
 1021 GTGATAAGGCTAACTACTCTACAACTCTCGATGCTTTTGGTATTATGGACTATACCT 1080
 1231 TAACTGGAAGTAGAGGATAATACGATGACACCAACCGTATCATACCGTTTATATGG 1290
 1081 TAACTGGAAGTAGAGGATAATACGATGACACCAACCGTATCATACCGTTTATATGG 1140
 1291 GCAAGCGACCCGAGAGAGATGCTAGCTATCATTTAGCCGCTGGTGCAGCGCAGC 1350
 1141 GCAAGCGACCCGAGAGAGATGCTAGCTATCATTTAGCTGCTGGTGGCCAGCGCAGC 1200
 1351 AAATGGTTC 1359
 1201 AGATTGTAC 1209

RESULT 5
 AAA37643
 ID AAA37643 standard; DNA; 2096 BP.
 XX
 AC AAA37643;
 XX
 DT 15-SEP-2003 (revised)
 DT 13-OCT-2000 (first entry)
 XX
 DE Chimeric SK-FBD coding sequence.
 XX
 KW Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
 KW plasminogen; human; fibronectin; thrombolytic therapy;
 KW cardiovascular disorder; ss.
 XX
 OS Streptococcus dysgalactiae subsp. equisimilis.
 OS Homo sapiens.
 OS Chimeric.
 XX
 PN EP1024192-A2.
 XX
 PD 02-AUG-2000.
 XX
 PF 23-DEC-1999; 99BP-00310541.
 XX
 PR 24-DEC-1998; 98IN-DE003825.
 XX
 XX (COUL) CSIR COUNCIL SCI IND RES.
 PA
 PI Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
 PI Yadav M;
 XX
 DR WPI; 2000-516032/47.
 XX
 XX Hybrid streptokinase-fibrin binding domain polypeptides useful for
 PT thrombolytic therapy comprises a streptokinase fused with fibrin binding
 PT domains of human fibronectin.
 XX
 PS Example 6; Fig 22b; 58pp; English.
 XX
 CC This sequence represents a chimeric streptokinase-fibrin binding domain
 CC (SK-FBD) protein coding sequence. The invention relates to a hybrid
 CC plasminogen activator (PA) comprises a polypeptide fusion between
 CC streptokinase (SK), which are capable of plasminogen (Pg) activation, and
 CC fibrin binding regions of human fibronectin, which are from fibrin
 CC binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA possesses the
 CC ability to bind with fibrin independently and also characteristically
 CC retains a PG activation ability which becomes evident only after a
 CC pronounced duration, or lag, after exposure of the PA to a suitable
 CC animal or human PG. The hybrid streptokinase-fibrin binding domain
 CC polypeptides are useful in thrombolytic therapy for various kinds of
 CC cardiovascular disorders. The hybrids have enhanced fibrin selectivity as
 CC well as kinetics of plasminogen activation that are distinct from that of
 CC natural streptokinase in being characterised by a temporary delay, or lag
 CC of several minutes in the natural rate of the catalytic conversion of

plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins
 CC can bind tightly with fibrin in blood clots soon after introduction into
 CC the vascular system without significantly activating the circulation of
 CC blood plasminogen to plasmin, thus aiding in the localisation of the
 CC plasminogen activation process to the site of pathological thrombus. This
 CC overcomes systemic plasminogen activation encountered during clinical use
 CC of streptokinase. (Updated on 15-SEP-2003 to standardise OS field)
 XX
 SQ Sequence 2096 BP; 643 A; 480 C; 484 G; 489 T; 0 U; 0 Other;
 Query Match 70.2%; Score 1165.8; DB 3; Length 2096;
 Best Local Similarity 99.4%; Pred. No. 0;
 Matches 1170; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 183 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACGCAATTCGTTGTT 242
 DB 588 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACGCAATTCGTTGTT 647
 QY 243 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 302
 DB 648 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 707
 QY 303 GATCTAACATCAGACCTGCTCATGAGGAGAAAGACAGAGCAAGCTTAAAGTCCAAATCA 362
 DB 708 GATCTAACATCAGACCTGCTCATGAGGAGAAAGACAGAGCAAGCTTAAAGTCCAAATCA 767
 QY 363 AAACCATTTGCTACTGATAGTGGCGGATGTCAATAACTTGAGAAAGCTGACTTACTA 422
 DB 768 AAACCATTTGCTACTGATAGTGGCGGATGTCAATAACTTGAGAAAGCTGACTTACTA 827
 QY 423 AAGGCTATTCAAGAAACAATTGATCGCTAACTCCACAGTAAACGACGACTACTTTGAGTGC 482
 DB 828 AAGGCTATTCAAGAAACAATTGATCGCTAACTCCACAGTAAACGACGACTACTTTGAGTGC 887
 QY 483 ATTGATTTTGCAGCGATGCAACCATTTACTGATCGAAACGCAAGGCTCTACTTTGCTGAC 542
 DB 888 ATTGATTTTGCAGCGATGCAACCATTTACTGATCGAAACGCAAGGCTCTACTTTGCTGAC 947
 QY 543 AAAGATGGTTCGGTAACCTTGGCCGACCAACCTGTCAGAGAAATTTTTCGTAAGCGGACAT 602
 DB 948 AAAGATGGTTCGGTAACCTTGGCCGACCAACCTGTCAGAGAAATTTTTCGTAAGCGGACAT 1007
 QY 603 GTGCGCGTTAGACCATATAAGAAACCAACCAATACAAACCAAGCGAAATCTGTTGATGTG 662
 DB 1008 GTGCGCGTTAGACCATATAAGAAACCAACCAATACAAACCAAGCGAAATCTGTTGATGTG 1067
 QY 663 GAATATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTTCAGACAGGCTCTCAA 722
 DB 1068 GAATATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTTCAGACAGGCTCTCAA 1127
 QY 723 GATATAGCTATTGAAAAACACTAGCTATCGGTGACACCATCAATCTCAAGAAATTAATA 782
 DB 1128 GATATAGCTATTGAAAAACACTAGCTATCGGTGACACCATCAATCTCAAGAAATTAATA 1187
 QY 783 GCTCAAGCACAAGCATTTTAAACAAACCAACCCAGCGCTATAGTATTATGAAAGCTGAC 842
 DB 1188 GCTCAAGCACAAGCATTTTAAACAAACCAACCCAGCGCTATAGTATTATGAAAGCTGAC 1247
 QY 843 TCCTCAATCTGCTACTCATGACAAATGACATTTTTCGCTAGCATTTTTCACCAATCGATCAAGAG 902
 DB 1248 TCCTCAATCTGCTACTCATGACAAATGACATTTTTCGCTAGCATTTTTCACCAATCGATCAAGAG 1307
 QY 903 TTTACTTACCGTGTATAAATTCGGGAAACAAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 962
 DB 1308 TTTACTTACCGTGTATAAATTCGGGAAACAAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 1367
 QY 963 AATGAAAGAAATAAACCAACACTGACCTGATCTCTGAGAAATATTAGCTCTTAAAAAGGG 1022
 DB 1368 AATGAAAGAAATAAACCAACACTGACCTGATCTCTGAGAAATATTAGCTCTTAAAAAGGG 1427
 QY 1023 GAAAGCGGTATGATCCCTTTTGTATGCGAGTCACTTGTAAACCTGTTTCAACCATCAATACGTT 1082
 DB 1428 GAAAGCGGTATGATCCCTTTTGTATGCGAGTCACTTGTAAACCTGTTTCAACCATCAATACGTT 1487

QY 1083 GATGTCGATACCAACGAATTGCTAAAAAGTGAGCAGCTCTTAACAGCTAGCGAAGCTAAC 1142
Dd |||||
Dd 1488 GATGTCGATACCAACGAATTGCTAAAAAGTGAGCAGCTCTTAACAGCTAGCGAAGCTAAC 1547
QY 1143 TTAGACTTCAGAGATTTATAGACTCCCTGCGATAGGCTAACTTACTTACACAAATCTC 1202
Dd |||||
Dd 1548 TTAGACTTCAGAGATTTATAGACTCCCTGCGATAGGCTAACTTACTTACACAAATCTC 1607
QY 1203 GATGCTTTTGGTATTATGGACTACTTAACTGCGAAAGTAGAGGATAATCAGATGAC 1262
Dd |||||
Dd 1608 GATGCTTTTGGTATTATGGACTACTTAACTGCGAAAGTAGAGGATAATCAGATGAC 1667
QY 1263 ACCAACCGTATACATACCGTTTATATGGCAAGCGACCCGAAGGAGAGAATGCTAGCTAT 1322
Dd |||||
Dd 1668 ACCAACCGTATACATACCGTTTATATGGCAAGCGACCCGAAGGAGAGAATGCTAGCTAT 1727
QY 1323 CATTTAGCGTGGTGGTGCAGGCGCAGCAAAATGGTTTC 1359
Dd |||||
Dd 1728 CATTTAGCTGGTGGTGGCCAGGCGCAGCAAGATTGTAC 1764

RESULT 6

AAA37642
ID AAA37642 standard; DNA; 1782 BP.
XX
AC AAA37642;
XX
DT 15-SEP-2003 (revised)
DT 13-OCT-2000 (first entry)
XX
DE Chimeric SK-FBD coding sequence.
XX
KW Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
KW plasminogen; human; fibronectin; thrombolytic therapy;
KW cardiovascular disorder; ss.
XX
OS Streptococcus dysgalactiae subsp. equisimilis.
OS Homo sapiens.
OS Chimeric.
XX
PN EP1024192-A2.
XX
PD 02-AUG-2000.
XX
PF 23-DEC-1999; 99EP-00310541.
XX
PR 24-DEC-1998; 98IN-DE003825.
XX
PA (COUL) CSIR COUNCIL SCI IND RES.
XX
PI Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
PI Yadav M;
XX
DR WPI; 2000-516032/47.
XX
PT Hybrid streptokinase-fibrin binding domain polypeptides useful for
PT thrombolytic therapy comprises a streptokinase fused with fibrin binding
PT domains of human fibronectin.
XX
PS Example 5; Fig 21b; 58pp; English.
XX
CC This sequence represents a chimeric streptokinase-fibrin binding domain
CC (SK-FBD) protein coding sequence. The invention relates to a hybrid
CC plasminogen activator (PA) comprises a polypeptide fusion between
CC streptokinase (SK), which are capable of plasminogen (PG) activation, and
CC fibrin binding regions of human fibronectin, which are from fibrin
CC binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA possesses the
CC ability to bind with fibrin independently and also characteristically
CC retains a PG activation ability which becomes evident only after a
CC pronounced duration, or lag, after exposure of the PA to a suitable
CC animal or human PG. The hybrid streptokinase-fibrin binding domain
CC polypeptides are useful in thrombolytic therapy for various kinds of

CC cardiovascular disorders. The hybrids have enhanced fibrin selectivity as
CC well as kinetics of plasminogen activation that are distinct from that of
CC natural streptokinase in being characterised by a temporary delay, or lag
CC of several minutes in the natural rate of the catalytic conversion of
CC plasminogen to plasmin (i.e. delayed-action thrombolytic). The proteins
CC can bind tightly with fibrin in blood clots soon after introduction into
CC the vascular system without significantly activating the circulating
CC blood plasminogen to plasmin, thus aiding in the localisation of the
CC plasminogen activation process to the site of pathological thrombus. This
CC overcomes systemic plasminogen activation encountered during clinical use
CC of streptokinase. (Updated on 15-SEP-2003 to standardise OS field)

XX Sequence 1782 BP; 573 A; 407 C; 377 G; 425 T; 0 U; 0 Other;

Query Match 69.2%; Score 1150.2; DB 3; Length 1782;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1152; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 183 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCAATTTGGTTGT 242
Dd |||||
Dd 538 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCAATTTGGTTGT 597
QY 243 AGCGTTGCTGTACTGTTTGAGGGACGAATCAAGACATTAGTCTTTAAATTTTGAATC 302
Dd |||||
Dd 598 AGCGTTGCTGTACTGTTTGAGGGACGAATCAAGACATTAGTCTTTAAATTTTGAATC 657
QY 303 GATCTAACATCAGACCTGCTCATGGAGGAAAGACAGAGCAAGGTTAAGTCCAAATCA 362
Dd |||||
Dd 658 GATCTAACATCAGACCTGCTCATGGAGGAAAGACAGAGCAAGGTTAAGTCCAAATCA 717
QY 363 AAACCATTTGCTACTGTAGTGGCGGATGTACATAAATTTGAGAAAGCTGACTTACTA 422
Dd |||||
Dd 718 AAACCATTTGCTACTGTAGTGGCGGATGTACATAAATTTGAGAAAGCTGACTTACTA 777
QY 423 AAGGCTATTCAAGAACCAATTTGATCGTTAAACGTCCACAGTACGACGACTCTTTGAGTGC 482
Dd |||||
Dd 778 AAGGCTATTCAAGAACCAATTTGATCGTTAAACGTCCACAGTACGACGACTCTTTGAGTGC 837
QY 483 ATTGATTTTGAAGCGATGCAACCATTTACTGATCGAAACGGCAAGGTTCTATTGCTGAC 542
Dd |||||
Dd 838 ATTGATTTTGAAGCGATGCAACCATTTACTGATCGAAACGGCAAGGTTCTATTGCTGAC 897
QY 543 AAAGATGTTGCGGTAACTTCCGCGACCCCACTGTCGAAGAAATTTTCTTAAGCGGACAT 602
Dd |||||
Dd 898 AAAGATGTTGCGGTAACTTCCGCGACCCCACTGTCGAAGAAATTTTCTTAAGCGGACAT 957
QY 603 GTGCGGTTAGACCATATAAGAAAAAACCAATATACAAACCAAGCGAAATCTGTTGATGTG 662
Dd |||||
Dd 958 GTGCGGTTAGACCATATAAGAAAAAACCAATATACAAACCAAGCGAAATCTGTTGATGTG 1017
QY 663 GAATATATCTGTACAGTTTACTTCCCTTAAACCTGATGACGATTTTCAAGACGGTCTCAA 722
Dd |||||
Dd 1018 GAATATATCTGTACAGTTTACTTCCCTTAAACCTGATGACGATTTTCAAGACGGTCTCAA 1077
QY 723 GATCTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTACTA 782
Dd |||||
Dd 1078 GATCTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTACTA 1137
QY 783 GCTCAAGCAAAAGCAATTTTAAACAAACCAACCCAGGCTATAGATTTATGAACGTCAC 842
Dd |||||
Dd 1138 GCTCAAGCAAAAGCAATTTTAAACAAACCAACCCAGGCTATAGATTTATGAACGTCAC 1197
QY 843 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 902
Dd |||||
Dd 1198 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 1257
QY 903 TTTACTTACCGTGTAAAAATCGGGAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 962
Dd |||||
Dd 1258 TTTACTTACCGTGTAAAAATCGGGAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 1317
QY 963 AATGAAGAAATAACACACATGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAGGG 1022
Dd |||||
Dd 1318 AATGAAGAAATAACACACATGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAGGG 1377

Qy	1023	GAAAGCCGTTATGATCCCTTTGATCGCAGTCACTTGAAACATGTTTCAACATCAAAATACGTT	1082
Db	1378	GAAAGCCGTTATGATCCCTTTGATCGCAGTCACTTGAAACATGTTTCAACATCAAAATACGTT	1437
Qy	1083	GATGTCGATACCAACGAATTCCTATAAAGTGAGCAGCTCTTTAACAGCTAGCGAACGTTAAC	1142
Db	1438	GATGTCGATACCAACGAATTCCTATAAAGTGAGCAGCTCTTTAACAGCTAGCGAACGTTAAC	1497
Qy	1143	TTAGACTTCAGAGATTTTATACGATCCTCGTGAATAAGGCTAAACTACCTCTCAACAATCTC	1202
Db	1498	TTAGACTTCAGAGATTTTATACGATCCTCGTGAATAAGGCTAAACTACCTCTCAACAATCTC	1557
Qy	1203	GATGCTTTTCGGTATTTATGCACTATACCTTAACTCGGAAAAGTAGAGGATAATCACGATGAC	1362
Db	1558	GATGCTTTTCGGTATTTATGCACTATACCTTAACTCGGAAAAGTAGAGGATAATCACGATGAC	1617
Qy	1263	ACCAACGGTATCATAAACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAAATGCTTAGCTAT	1322
Db	1618	ACCAACGGTATCATAAACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAAATGCTTAGCTAT	1677
Qy	1323	CATTTAGCCGGTGGT	1337
Db	1678	CATTTAGCCCTATGAT	1692

RESULT 7

AAx80497	
ID	AAx80497 standard; cDNA; 2385 BP.
XX	
XX	
XX	AAx80497;
XX	
DT	26-AUG-1999 (first entry)
XX	
DE	Streptokinase and maltose binding protein fusion protein encoding cDNA.
XX	
KW	Streptococcus; streptokinase; fibrin-dependent plasminogen activator;
KW	nSk; rSk; bacterial; blood clot; thrombotic condition;
KW	myocardial infarction; venous thrombosis; pulmonary embolism;
KW	cerebral thrombosis; graft thrombosis; arterial thrombosis; ss.
XX	
OS	Streptococcus dysgalactiae subsp. equisimilis.
OS	Synthetic.
XX	
PN	WO9931247-A1.
XX	
PD	24-JUN-1999.
XX	
PF	15-DEC-1998; 98WO-US026694.
XX	
PR	15-DEC-1997; 97US-0069497P.

XX Reed GL;
 XX
 XX WPI; 1999-395183/33.
 DR P-PSDE; AAY24797.
 XX
 XX N-terminally deleted streptokinase.
 PT
 XX
 XX Example; Page 45-48; 73pp; English.
 PS
 XX
 XX The present invention describes an isolated bacterial protein that
 CC induces fibrin-dependent plasminogen activation in a pharmaceutical
 CC composition for dissolving blood clots. Also described are: (1) a
 CC composition comprising an isolated modified streptokinase, the
 CC modification being removal of amino acid residues in the amino terminus;
 CC (2) a method for dissolving a blood clot in a subject, comprising
 CC administering to the subject a fibrin-dependent streptokinase protein; a
 CC nucleic acid (1) encoding a modified bacterial streptokinase; (3) an
 CC expression vector comprising (1); and (4) a host cell transformed with
 CC the expression vector of (3). The pharmaceutical composition comprising a

bacterial fibrin-dependent plasminogen activator is useful for dissolving blood clots in patients with a thrombotic condition, e.g. myocardial infarction, venous thrombosis, pulmonary embolism, cerebral thrombosis, graft thrombosis and arterial thrombosis. The modified streptokinase can also be used in non-human mammals. Streptokinase activation of plasminogen is at least 10-fold, preferably 100-fold greater in the presence of fibrin than in the absence of fibrin. The modified streptokinase has at least one amino acid substitution that inactivates a substrate site for proteolytic cleavage. This reduces the rate of degradation of the streptokinase at least two-fold. The present sequence encodes a streptokinase and maltose binding protein fusion protein from an example of the present invention

Sequence 2385 BP; 745 A; 547 C; 549 G; 544 T; 0 U; 0 Other;

Query Match	69.2%	Score 1150;	DB 2;	Length 2385;
Best Local Similarity	99.6%	Pred. No. 0;		

180 ATGATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAAATTGGTT 239

1141 AGGATTGGACCTGAGTGGCTAGACCGTCCATCTGTCAACACAGCCAAATTAGTTT 1200

200

[illegible]

50 1281 ATCGATCTAAGCATCAGGACCTGCTCATGTGGAGGAAAGAGACAGAGCAAGGCCTAAGTCCAGAA 1320

360 TCAAAACCATTTGCTACTGATGATGGCGGATGTCACATAAACTTGAGAAAGCTGACTTAA 419

Db 1321 TCAAAACCATTTGCTACTGATAGTGGCGCGATGTCACTAAACTTNGAGAAAGCTGACTTA 1380

QY 420 CTAAAGGCTATTCAAGAAACAATTGATCGCTAACGTCCACAGTAAACGACACTACTTTTGAG 479

Db 1381 CTAAAGGCTATTCAAGAACAAATTGATCGCTAACGTCCACAGTAACGACGACTACTTTTGAG 1440

QV 480 GTCA TTGAT TTTC AAGCGATGCAACCATTA CTGATCGAAACGGCAAGTCTACTTTGCT 539

D_b 1441 GTCAATTGATTTTGTCAAGCGATGCAACCATTA CTGATCCGAAACGGCAAGTCTACATTGCT 1500

540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000

[illegible]

	TOTAL	MALE	FEMALE
0			
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			
16			
17			
18			
19			
20			
21			
22			
23			
24			
25			
26			
27			
28			
29			
30			
31			
32			
33			
34			
35			
36			
37			
38			
39			
40			
41			
42			
43			
44			
45			
46			
47			
48			
49			
50			
51			
52			
53			
54			
55			
56			
57			
58			
59			
60			
61			
62			
63			
64			
65			
66			
67			
68			
69			
70			
71			
72			
73			
74			
75			
76			
77			
78			
79			
80			
81			
82			
83			
84			
85			
86			
87			
88			
89			
90			
91			
92			
93			
94			
95			
96			
97			
98			
99			
100			

QY 660 GTGGAAATATACIGTATCAGTTTACTCCCTTAAACCCCTGAIGACGATTTGAGACCAGGCTC 719

Db 1621 GTGGAATATACTGTACAGTTACTCCCTTAAACCCCTGATGACGATTCAGACCAGGTC TC 1680

Qy 720 AAAGATACTAAGCTATTGAAACAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATTA 779

Db 1681 AAAGATACTAAGCTATTGAAAACACTAGCTATCGGTGACCATCATCTCAAGAAATTA 1740

QY 780 CTAGCTCAGCACAAAGCATTTTAAACAAAACCCAGGCTATACGATTTATGAACGT 839

db 1741 CTAGCTCAAGCACAAGCATTTTAAACAAAACCACCCAGGCTATACGATTATGAACGT 1800

840 GAGTCCCTCAATTCCCTCACTTCATGACACATGTCATGTTTCCGTAAGATGGATCAA 899

1000

[illegible]

QY CIGAAIGAGGAAAATAAACAACATCGACCCTGGATCTCTGGAGAAATAATTACGTCCTTAAAGAAAA TGTG

Db 1921 CTGAATGAAGAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAA 1980
Qy 1020 GGGGAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTTGAAACTGTTTCAACCATCAAAATAC 1079
Db 1981 GGGGAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTTGAAACTGTTTCAACCATCAAAATAC 2040
Qy 1080 GTTGATGTCGATACCAACGAATTGCTTAAAGTGAGCAGCTCTTTAAACAGCTAGCGAACGT 1139
Db 2041 GTTGATGTCGATACCAACGAATTGCTTAAAGTGAGCAGCTCTTTAAACAGCTAGCGAACGT 2100
Qy 1140 AACTTAGACTTCAGAGATTTATACGATCCCTGCTGATAGGCTAACTACTCTACAACAT 1199
Db 2101 AACTTAGACTTCAGAGATTTATACGATCCCTGCTGATAGGCTAACTACTCTACAACAT 2160
Qy 1200 CTCGATGCTTTTGGTATTATGGACTATACCTTAACTGGAAGAGTAGAGGATAATCACAT 1259
Db 2161 CTCGATGCTTTTGGTATTATGGACTATACCTTAACTGGAAGAGTAGAGGATAATCACAT 2220
Qy 1260 GACACCAACGGTATCATACCGTTTATATGGGCAAGCGACCCGAAAGGAGAGAATGCTAGC 1319
Db 2221 GACACCAACGGTATCATACCGTTTATATGGGCAAGCGACCCGAAAGGAGAGAATGCTAGC 2280
Qy 1320 TATCATTTAGCGGTGGT 1337
Db 2281 TATCATTTAGCGCTATGAT 2298

RESULT 8
AAQ11651
ID AAQ11651 standard; DNA; 2030 BP.
XX
AC AAQ11651;
XX
DT 08-JUL-1991 (first entry)
XX
DE PB-FB-SK fusion construct.
XX
KW Fibrin-binding protein; fibrinolysis; intravascular thrombi; fibrinogen;
KW streptokinase; fusion protein; 58.
XX
OS Staphylococcus aureus.
XX
FH Key
CDS 10..184
FT /tag= a
FT /label= FB monomer
FT 185..358
CDS
FT /tag= b
FT /label= FB monomer
FT 359..1601
CDS
FT /tag= c
FT /label= streptokinase
XX
PN US5011686-A.
XX
PD 30-APR-1991.
XX
PF 15-NOV-1989; 89US-00437769.
XX
PR 21-SEP-1987; 87US-00099242.
XX
PA (CREA-) CREATIVE BIOMOLEC.
XX
PI Pang RHL;
XX
DR WPI; 1991-140198/19.
DR P-PSDB; AAR11829.
XX
XX Imparting injectable fibrinolytic agent - with affinity for intravascular
PT thrombus, by linking agent to fibrin binding domain.
XX
PS Disclosure; Fig 5; 18pp; English.

CC The DNA encodes an FB-FB dimer linked to the streptokinase coding
CC sequence. The FB fragment has selective affinity for fibrin, low affinity
CC for fibrinogen, and minimal immunogenicity, imparting thrombus-targeting
CC capability. See also AAQ11649 and AAQ11650
XX
SQ Sequence 2030 BP; 673 A; 464 C; 406 G; 487 T; 0 U; 0 Other;

Query Match 69.2%; Score 1148.8; DB 2; Length 2030;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1153; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 178 COATGATTCCTGGACCTCAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCCAATTTGG 237
Db 353 CGAAGATTGCTGGACCTCAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCCAATTTAG 412
Qy 238 TTGTTAGCGTTGCTGGTACTGTTGAGGGGAGCAATCAAGACATTAAGTCTTAAATTTTGG 297
Db 413 TTGTTAGCGTTGCTGGTACTGTTGAGGGGAGCAATCAAGACATTAAGTCTTAAATTTTGG 472
Qy 298 AAATCGATCTAAACATCACGACCTGCTCATGAGGAAAGACAGAGCAAGGCTTTAAGTCCAA 357
Db 473 AAATCGATCTAAACATCACGACCTGCTCATGAGGAAAGACAGAGCAAGGCTTTAAGTCCAA 532
Qy 358 AATCAAAACCAATTTGCTACTGATAGTGGCGGATGTCAATAAACTTGAGAAAGCTGACT 417
Db 533 AATCAAAACCAATTTGCTACTGATAGTGGCGGATGTCAATAAACTTGAGAAAGCTGACT 592
Qy 418 TACTAAAGGCTATTCAAGAACAAATTTGATCGCTTAAGTCCACAGTAACGACGACTACTTTG 477
Db 593 TACTAAAGGCTATTCAAGAACAAATTTGATCGCTTAAGTCCACAGTAACGACGACTACTTTG 652
Qy 478 AGGTCATTGATTTTGCAGGCGATGCAACCATTTACTGATCGAAGCGGAGGCTACTTTTG 537
Db 653 AGGTCATTGATTTTGCAGGCGATGCAACCATTTACTGATCGAAGCGGAGGCTACTTTTG 712
Qy 538 CTGACAAAGATGTTTGGTAACTTTGCCGACCAACCTGTCCAAAGAAATTTTGTCTAAGCG 597
Db 713 CTGACAAAGATGTTTGGTAACTTTGCCGACCAACCTGTCCAAAGAAATTTTGTCTAAGCG 772
Qy 598 GACATGTCGCGTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTTG 657
Db 773 GACATGTCGCGTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTTG 832
Qy 658 ATGTGGAATATACCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTACAGCAGTC 717
Db 833 ATGTGGAATATACCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTACAGCAGTC 892
Qy 718 TCAAGATATCTAAGCTATTGAAACACATAGCTATCGGTGACACCATCATCTCAAGAAAT 777
Db 893 TCAAGATATCTAAGCTATTGAAACACATAGCTATCGGTGACACCATCATCTCAAGAAAT 952
Qy 778 TACTAGCTCAAGCACAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTATGAAC 837
Db 953 TACTAGCTCAAGCACAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTATGAAC 1012
Qy 838 GTGACTCTCAATCGTCACTCATGCAATGACATTTTCCGTACGATTTTACCAATGGATC 897
Db 1013 GTGACTCTCAATCGTCACTCATGCAATGACATTTTCCGTACGATTTTACCAATGGATC 1072
Qy 898 AAGAGTTTACTTACCGGTGTTAAATAATCGGGAACAGCTTATAGGATCAATAAAAAATCTG 957
Db 1073 AAGAGTTTACTTACCGGTGTTAAATAATCGGGAACAGCTTATAGGATCAATAAAAAATCTG 1132
Qy 958 GTCTGAATGAAGAAATAAACAACACTGACCTGTCTCTGAGAAATATTACGCTCTTAAAA 1017
Db 1133 GTCTGAATGAAGAAATAAACAACACTGACCTGTCTCTGAGAAATATTACGCTCTTAAAA 1192
Qy 1018 AAGGGGAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTTCAACATCAAT 1077
Db 1193 AAGGGGAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTTCAACATCAAT 1252
Qy 1078 ACGTTGATGTCGATACCAACGAATTTGCTTAAAGTGAGCAGCTCTTAAACAGCTAGCGAAC 1137

Db 1253 ACGTTGATCGCATACCAACGAATTGCTGTAATAAGTGAGCAGCTCTTTAAACAGCTAGCGAAC 1312
 Qy 1138 GTAACCTAGACTTCAGAGATTTATACGATCCTCGTGATAGGCTAAACACTACTCTACAAACA 1197
 Db 1313 GFACTTAGACTTCAGAGATTTATACGATCCTCGTGATAGGCTAAACACTACTCTACAAACA 1372
 Qy 1198 AFTCTCGATGCTTTGGTATTATGGAATACTTAACCTTAACCTGGAAGAGTAGAGATAATCACG 1257
 Db 1373 AFTCTCGATGCTTTGGTATTATGGAATACTTAACCTTAACCTGGAAGAGTAGAGATAATCACG 1432
 Qy 1258 ATGACACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGGAAGGAGAGAAATGCTTA 1317
 Db 1433 ATGACACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGGAAGGAGAGAAATGCTTA 1492
 Qy 1318 GCTATCATTTAGCCGGTGGT 1337
 Db 1493 GCTATCATTTAGCCTATGAT 1512

RESULT 9

AXX80492
 ID AAX80492 standard; cDNA; 1242 BP.
 XX AC

AXX80492;

17-OCT-2003 (revised)
 26-AUG-1999 (first entry)

Streptococcus equisimilis native streptokinase encoding cDNA.

Streptococcus; streptokinase; fibrin-dependent plasminogen activator;
 KW rSK; bacterial; blood clot; thrombotic condition;
 KW myocardial infarction; venous thrombosis; pulmonary embolism;
 KW cerebral thrombosis; graft thrombosis; arterial thrombosis; ss.

Streptococcus dysgalactiae subsp. equisimilis.

WO9931247-A1.

24-JUN-1999.

15-DEC-1998; 98WO-US026694.

15-DEC-1997; 97US-0069497P.

(HARD) HARVARD COLLEGE.

Reed GL;

WPI; 1999-395183/33.

P-PSDB; AAY24794.

N-terminally deleted streptokinase.

Claim 44; Page 58-60; 73pp; English.

The present invention describes an isolated bacterial protein that induces fibrin-dependent plasminogen activation in a pharmaceutical composition for dissolving blood clots. Also described are: (1) a composition comprising an isolated modified streptokinase, the modification being removal of amino acid residues in the amino terminus; (2) a method for dissolving a blood clot in a subject, comprising administering to the subject a fibrin-dependent streptokinase protein; a nucleic acid (1) encoding a modified bacterial streptokinase; (3) an expression vector comprising (1); and (4) a host cell transformed with the expression vector of (3). The pharmaceutical composition comprising a bacterial fibrin-dependent plasminogen activator is useful for dissolving blood clots in patients with a thrombotic condition, e.g. myocardial infarction, venous thrombosis, pulmonary embolism, cerebral thrombosis, graft thrombosis and arterial thrombosis. The modified streptokinase can also be used in non-human mammals. Streptokinase activation of plasminogen is at least 10-fold, preferably 100-fold greater in the presence of fibrin than in the absence of fibrin. The modified

CC streptokinase has at least one amino acid substitution that inactivates a substrate site for proteolytic cleavage. This reduces the rate of degradation of the streptokinase at least two-fold. The present sequence encodes native streptokinase (nsk). (Updated on 17-OCT-2003 to standardise OS field)

XX SQ Sequence 1242 BP; 424 A; 267 C; 237 G; 314 T; 0 U; 0 Other;

Query Match 69.28; Score 1148.6; DB 2; Length 1242;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 1151; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 183 ATTGCTCGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCCAAATGTTGTT 242
 Db 1 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCCAAATGTTGTT 60
 Qy 243 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAATTTTGAATC 302
 Db 61 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAATTTTGAATC 120
 Qy 303 GATCTAATCATCAGACCTGCTCATGGAGGAAAGACAGACGAGGCTTAAGTCCAAATCA 362
 Db 121 GATCTAATCATCAGACCTGCTCATGGAGGAAAGACAGACGAGGCTTAAGTCCAAATCA 180
 Qy 363 AAACCATTTGCTACTGATAGTGGCGCATGTCATATAAATTGAGAAAGCTGACTTACTA 422
 Db 181 AAACCATTTGCTACTGATAGTGGCGCATGTCATATAAATTGAGAAAGCTGACTTACTA 240
 Qy 423 AAGGCTATTCAAGAACAAATTGATCGTTAAGCTCCACAGTAACGACGACTACTTTGAGTGC 482
 Db 241 AAGGCTATTCAAGAACAAATTGATCGTTAAGCTCCACAGTAACGACGACTACTTTGAGTGC 300
 Qy 483 ATTGATTTTGCAGCGATGCAACCATTAAGTATCGAATCGCAAGCGCAAGGCTACTTTGCTGAC 542
 Db 301 ATTGATTTTGCAGCGATGCAACCATTAAGTATCGAATCGCAAGCGCAAGGCTACTTTGCTGAC 360
 Qy 543 AAAGATGTTTCGGTAACTTTGCCGACCCAACTCTCCAAAGAAATTTTGTAAAGCGGACAT 602
 Db 361 AAAGATGTTTCGGTAACTTTGCCGACCCAACTCTCCAAAGAAATTTTGTAAAGCGGACAT 420
 Qy 603 GTGCGGTTAGACCATATAAAGAAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 662
 Db 421 GTGCGGTTAGACCATATAAAGAAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 480
 Qy 663 GAATATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAA 722
 Db 481 GAATATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAA 540
 Qy 723 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATTA 782
 Db 541 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATTA 600
 Qy 783 GCTCAAGCACAAGCAATTTTAAACAAAAACCAAGGCTATACGATTTATGAACGTCAC 842
 Db 601 GCTCAAGCACAAGCAATTTTAAACAAAAACCAAGGCTATACGATTTATGAACGTCAC 660
 Qy 843 TCCTCAATCGTCACTCAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 902
 Db 661 TCCTCAATCGTCACTCAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 720
 Qy 903 TTTTACTTACCGTGTAAAAATCGGAAACAGCTTATAGGATCAATAAAAAATCTGCTCG 962
 Db 721 TTTTACTTACCGTGTAAAAATCGGAAACAGCTTATAGGATCAATAAAAAATCTGCTCG 780
 Qy 963 AATGAAGAAATAAACAACACTGACTGATCTCTGAGAAATATTAAGTCTTAAAAAAGGG 1022
 Db 781 AATGAAGAAATAAACAACACTGACTGATCTCTGAGAAATATTAAGTCTTAAAAAAGGG 840
 Qy 1023 GAAAGCCGATGATCCCTTTGATCGCAGTCACTTTGAAACCTTTCAACATCAATACGTT 1082
 Db 841 GAAAGCCGATGATCCCTTTGATCGCAGTCACTTTGAAACCTTTCAACATCAATACGTT 900
 Qy 1083 GATGTCGATACCAAGAAATTGCTTAAAAAGTGAGCAGCTCTTTAACAGCTAGCGAAGTAA 1142

Db 781 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 840
 Qy 1023 GAAGAGCCGTATGATCCCTTTGATGCGAGTCACCTTGAAACTGTTCCACATCAATACGTT 1082
 Db 841 GAAAGCCGTATGATCCCTTTGATGCGAGTCACCTTGAAACTGTTCCACATCAATACGTT 900
 Qy 1083 GATGTCGATACCAACAAATTTGCTTAAAGGTGACGAGCTCTTAAACAGCTAGCGAACGTAAC 1142
 Db 901 GATGTCGATACCAACAAATTTGCTTAAAGGTGACGAGCTCTTAAACAGCTAGCGAACGTAAC 960
 Qy 1143 TTAGACTTCAGAGATTATACGATCCTCGTGATAAGGCTAAACTACTCTACAAATCTC 1202
 Db 961 TTAGACTTCAGAGATTATACGATCCTCGTGATAAGGCTAAACTACTCTACAAATCTC 1020
 Qy 1203 GATGCTTTTGGTATTATGACTATACCTTAACCTGGAAGGTAGAGGATATCAGGATGAC 1262
 Db 1021 GATGCTTTTGGTATTATGACTATACCTTAACCTGGAAGGTAGAGGATATCAGGATGAC 1080
 Qy 1263 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGGAAGGAGAGAAATGCTAGCTAT 1322
 Db 1081 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGGAAGGAGAGAAATGCTAGCTAT 1140
 Qy 1323 CATTTAGCCGCGGT 1337
 Db 1141 CATTTAGCCGTATGAT 1155

RESULT 11
 ID ABA05546
 XX ABA05546 standard; cDNA; 1254 BP.
 XX AC ABA05546;
 XX DT 26-FEB-2002 (first entry)
 XX DE Streptokinase cDNA.
 XX KW Streptokinase; cerebroprotective; cardiact; gene therapy; fusion;
 XX gene therapy; thrombolytic; angina; myocardial infarction; stroke;
 XX Unidentified.
 XX WO200185100-A2.
 XX PD 15-NOV-2001.
 XX PF 10-MAY-2001; 2001WO-US015209.
 XX PR 11-MAY-2000; 2000US-00569920.
 XX (GEO) GEN HOSPITAL CORP.
 XX Reddy VB, Lerner E;
 XX WPI; 2002-062184/08.
 XX New fusion protein or conjugate, useful for treating unstable angina,
 XX acute myocardial infarction or stroke, comprises a vasodilator
 XX polypeptide and a thrombolytic polypeptide, or active fragments of the
 XX polypeptides.
 XX Example 1; Fig 2; 37pp; English.
 XX The invention relates to a fusion protein or a conjugate comprising a
 XX vasodilator polypeptide, or its active fragment, and a thrombolytic
 XX polypeptide or its active fragment. The protein is useful for treating a
 XX subject suffering from a partially or totally occluded blood vessel,
 XX causing unstable angina, acute myocardial infarction or stroke. The
 XX polynucleotide encoding the polypeptide is useful in gene therapy. The
 XX vasodilatory action of the protein allows for the use of lower doses of a
 XX thrombolytic while maintaining the clot dissolving effectiveness of the

thrombolytic, and the use of lower doses of the thrombolytic reduces
 associated side effects. The present sequence is the streptokinase cDNA
 used in the construction of a Maxadilan-Streptokinase fusion protein.
 Maxadilan is a vasodilator peptide produced by the salivary gland of the
 New World sand fly
 Query Match 69.2%; Score 1148.6; DB 6; Length 1254;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 1151; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Qy 183 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCAATTTGGTCTT 242
 Db 7 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCAATTTGGTCTT 66
 Qy 243 AGCGTTCTGCTACTGTGTAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 302
 Db 67 AGCGTTCTGCTACTGTGTAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 126
 Qy 303 GATCTAATCATCAGCACTGCTCATGGAGGAAAGACAGCAAGGCTTAAGTCCAAATCA 362
 Db 127 GATCTAATCATCAGCACTGCTCATGGAGGAAAGACAGCAAGGCTTAAGTCCAAATCA 186
 Qy 363 AAACCATTTGCTACTGATAGTGGCGCGATGTACATAAACTTGAGAAAGCTGACTTACTA 422
 Db 187 AAACCATTTGCTACTGATAGTGGCGCGATGTACATAAACTTGAGAAAGCTGACTTACTA 246
 Qy 423 AAGGCTATTCAAGAACAAATTTGATCGCTAAAGTCCACAGTAAACGACGACTTCTTTGAGTGC 482
 Db 247 AAGGCTATTCAAGAACAAATTTGATCGCTAAAGTCCACAGTAAACGACGACTTCTTTGAGTGC 306
 Qy 483 ATTGATTTTGCACGAGTGCACCAATTTACTGATCGAAACGCAAGGCTTACTTTGCTGAC 542
 Db 307 ATTGATTTTGCACGAGTGCACCAATTTACTGATCGAAACGCAAGGCTTACTTTGCTGAC 366
 Qy 543 AAGATGTTTTCGGTAAACCTTTCGCCACCCCACTCTGCAAGAAATTTTGTCTAAGCGGACAT 602
 Db 367 AAGATGTTTTCGGTAAACCTTTCGCCACCCCACTCTGCAAGAAATTTTGTCTAAGCGGACAT 426
 Qy 603 GTGCGCGTTAGACCATATATAAGAAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 662
 Db 427 GTGCGCGTTAGACCATATATAAGAAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 486
 Qy 663 GAATATATCTGTAAGTTTACTTCCCTTAAACCCCTGATGACGATTTTTCAGACGAGTCTCAA 722
 Db 487 GAATATATCTGTAAGTTTACTTCCCTTAAACCCCTGATGACGATTTTTCAGACGAGTCTCAA 546
 Qy 723 GATCTAAGCTATTGAAACACACTAGCTATCGGTGACACCATCATCTCTCAAGAAATTACTA 782
 Db 547 GATCTAAGCTATTGAAACACACTAGCTATCGGTGACACCATCATCTCTCAAGAAATTACTA 606
 Qy 783 GCTCAAGCAACAAGCAATTTTAAACCAAAACCAAGCGCTATACGATTTTATGAAAGTGCAC 842
 Db 607 GCTCAAGCAACAAGCAATTTTAAACCAAAACCAAGCGCTATACGATTTTATGAAAGTGCAC 666
 Qy 843 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 902
 Db 667 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 726
 Qy 903 TTTTACTTACCGTGTAAAGATCGGGAACAGCTTATAGGATCAATAAAATCTGGTCTG 962
 Db 727 TTTTACTTACCGTGTAAAGATCGGGAACAGCTTATAGGATCAATAAAATCTGGTCTG 786
 Qy 963 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1022
 Db 787 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 846
 Qy 1023 GAAAGCCGTATGATCCCTTTGATGCGAGTCACCTTGAAACTGTTCCACATCAATACGTT 1082
 Db 847 GAAAGCCGTATGATCCCTTTGATGCGAGTCACCTTGAAACTGTTCCACATCAATACGTT 906
 Qy 1083 GATGTCGATACCAACAAATTTGCTTAAAGGTGACGAGCTCTTAAACAGCTAGCGAACGTAAC 1142

Db 907 GATGTCATACCAAGCAATTCCTAAAGAGTGAGCAGCTCTTAACAGCTAGCGAACGTAAAC 966
 QY 1143 TTAGACTTCAGAGATTTATAGGATCCTCGTATAGGCTAAATCTCTACAAATCTC 1202
 Db 967 TTAGACTTCAGAGATTTATAGGATCCTCGTATAGGCTAAATCTCTACAAATCTC 1026
 QY 1203 GATGCTTTGGTATATAGGACTATACCTTAATCTGAAAGTAGAGGATAATCAGATGAC 1262
 Db 1027 GATGCTTTGGTATATAGGACTATACCTTAATCTGAAAGTAGAGGATAATCAGATGAC 1086
 QY 1263 ACCAACCGTATCAACCGTTTATATGGGCAAGCGACCCGAGGAGAGAGAAATCTAGCTAT 1322
 Db 1087 ACCAACCGTATCAACCGTTTATATGGGCAAGCGACCCGAGGAGAGAGAAATCTAGCTAT 1146
 QY 1323 CATTTAGCCGGTGGT 1337
 Db 1147 CATTTAGCCCTATGAT 1161

RESULT 12

ABA05547
 ID ABA05547 standard; DNA; 8893 BP.

XX AC ABA05547;

XX DT 26-FEB-2002 (first entry)

XX DE Maxadilan-streptokinase fusion protein plasmid pTVB3maxstk.

XX KW Streptokinase; cerebroprotective; cardiant; gene therapy; fusion;

XX KW vasodilator; thrombolytic; angina; myocardial infarction; stroke;

XX KW gene therapy; maxadilan; sand fly; plasmid; ds.

XX OS Lutomyia longipalpis.

XX OS Unidentified.

XX OS Synthetic.

XX OS Chimeric.

XX PN WO200185100-A2.

XX PD 15-NOV-2001.

XX PF 10-MAY-2001; 2001WO-US015209.

XX PR 11-MAY-2000; 2000US-00569920.

XX PA (GEO) GEN HOSPITAL CORP.

XX PI Reddy VB, Lerner E;

XX PS WPI; 2002-062184/08.

XX PT New fusion protein or conjugate, useful for treating unstable angina,

XX PT acute myocardial infarction or stroke, comprises a vasodilator

XX PT polypeptide and a thrombolytic polypeptide, or active fragments of the

XX PT polypeptides.

XX PS Example 1; Fig 3; 37pp; English.

XX CC The invention relates to a fusion protein or a conjugate comprising a

XX CC vasodilator polypeptide, or its active fragment, and a thrombolytic

XX CC polypeptide or its active fragment. The protein is useful for treating a

XX CC subject suffering from a partially or totally occluded blood vessel,

XX CC causing unstable angina, acute myocardial infarction or stroke. The

XX CC polynucleotide encoding the polypeptide is useful in gene therapy. The

XX CC vasodilatory action of the protein allows for the use of lower doses of a

XX CC thrombolytic while maintaining the clot dissolving effectiveness of the

XX CC thrombolytic, and the use of lower doses of the thrombolytic reduces

XX CC associated side effects. The present sequence is the plasmid used for the

XX CC preparation of Maxadilan-Streptokinase fusion protein. It contains

XX CC maxadilan cDNA, which encodes a vasodilator peptide produced by the

XX CC salivary gland of the New World sand fly, and streptokinase cDNA

XX

SQ Sequence 8893 BP; 2366 A; 2106 C; 2246 G; 2175 T; 0 U; 0 Other;

Query Match

Best Local Similarity 69.2%; Score 1148.6; DB 6; Length 8893;

Matches 1151; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY

183 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAACAGCAATTTGGTTGTT 242

Db

5935 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAACAGCAATTTGGTTGTT 5994

QY

243 AGCGTTGCTGTGTAATCTGTTGAGGGGCGAATCAAGACATTTAGTCTTTAAATTTTGAATTC 302

Db

5995 AGCGTTGCTGTGTAATCTGTTGAGGGGCGAATCAAGACATTTAGTCTTTAAATTTTGAATTC 6054

QY

303 GATCTAAACATCACGACCTGCTCATGAGGAAAGACAGCAAGGCTTAAGTCCAAATCA 362

Db

6055 GATCTAAACATCACGACCTGCTCATGAGGAAAGACAGCAAGGCTTAAGTCCAAATCA 6114

QY

363 AAACCATTTGCTACTGATAGTGGCGCGATGTCAATAACTTGAGAAAGCTGACTTACTTA 422

Db

6115 AAACCATTTGCTACTGATAGTGGCGCGATGTCAATAACTTGAGAAAGCTGACTTACTTA 6174

QY

423 AAGGCTATTCAAGAAACAAATTGATCGCTAACCGTCCACAGTAACGACGACTTCTTGGAGTTC 482

Db

6175 AAGGCTATTCAAGAAACAAATTGATCGCTAACCGTCCACAGTAACGACGACTTCTTGGAGTTC 6234

QY

483 ATTGATTTTGCAGCGATGCAACCAATTACTGATCGCTAACCGTCCACAGTAACGACGACTTCTTGGAGTTC 542

Db

6235 ATTGATTTTGCAGCGATGCAACCAATTACTGATCGCTAACCGTCCACAGTAACGACGACTTCTTGGAGTTC 6294

QY

543 AAAGATGTTTCGGTAACTTCCGACCCCAACCTGTCGAAGAAATTTTTCGTAAGCGGACAT 602

Db

6295 AAAGATGTTTCGGTAACTTCCGACCCCAACCTGTCGAAGAAATTTTTCGTAAGCGGACAT 6354

QY

603 GTGCGCGTTAGACCATATAAAGAAAAACCAATACAAAAACCAAGCGGAAATCTGTTGATGTCG 662

Db

6355 GTGCGCGTTAGACCATATAAAGAAAAACCAATACAAAAACCAAGCGGAAATCTGTTGATGTCG 6414

QY

663 GAATATCTGTACAGTTTATCTCCCTTAAACCTGATGACGATTTTCAGACCAAGGCTCTCAAA 722

Db

6415 GAATATCTGTACAGTTTATCTCCCTTAAACCTGATGACGATTTTCAGACCAAGGCTCTCAAA 6474

QY

723 GATCTAAGCTATTGAAACACACTAGCTATCGGTGACACATCAGATCTCAAGAAATTACTA 782

Db

6475 GATCTAAGCTATTGAAACACACTAGCTATCGGTGACACATCAGATCTCAAGAAATTACTA 6534

QY

783 GCTCAAGCACAAAGCATTTTAAACAAAAACCAACCCAGGCTATACGATTTATGAACGTCAC 842

Db

6535 GCTCAAGCACAAAGCATTTTAAACAAAAACCAACCCAGGCTATACGATTTATGAACGTCAC 6594

QY

843 TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 902

Db

6595 TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 6654

QY

903 TTTTACTTCGTGTTAAAAATCGGGAAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 962

Db

6655 TTTTACTTCGTGTTAAAAATCGGGAAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 6714

QY

963 AATCAAGAAAAATAAACACACTGACCTGATCTCTGGAATATATTACGCTCTTAAAAAAGGG 1022

Db

6715 AATCAAGAAAAATAAACACACTGACCTGATCTCTGGAATATATTACGCTCTTAAAAAAGGG 6774

QY

1023 GAAAGCGGTATGATCCCTTTGATCGGAGTCACCTTGAACCTGTTCAACCATCAAAATACGTT 1082

Db

6775 GAAAGCGGTATGATCCCTTTGATCGGAGTCACCTTGAACCTGTTCAACCATCAAAATACGTT 6834

QY

1083 GATGTCGATACCAACGAATTTGCTAAAAAGTGAAGAGCTCTTTAAACAGCTAGCGAAGCTAAC 1142

Db

6835 GATGTCGATACCAACGAATTTGCTAAAAAGTGAAGAGCTCTTTAAACAGCTAGCGAAGCTAAC 6894

QY

1143 TTAGACTTCAGAGATTTATAGGATCCTCGTGTATAGGCTAACTTCTTACACAAATCTC 1202

Db 6895 TTAGACTTCAGAGATTATATACGATCCCTCGTATAAGGCTAAACTACTCTACAAACAATCTC 6954
 QY 1203 GATGCTTTTGGTATTATGGACTACTTAACTGGAAGTAGAGGATAATCAGATGAC 1262
 Db 6955 GATGCTTTTGGTATTATGGACTACTTAACTGGAAGTAGAGGATAATCAGATGAC 7014
 QY 1263 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAATGCTAGCTAT 1322
 Db 7015 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAATGCTAGCTAT 7074
 QY 1323 CATTTAGCCGGTGGT 1337
 Db 7075 CATTTAGCCCTATGAT 7089

RESULT 13
 AAX16632
 ID AAX16632 standard; DNA; 1242 BP.
 AC AAX16632;
 XX
 DT 17-OCT-2003 (revised)
 DT 04-MAY-1999 (first entry)
 XX
 DE Streptococcus equisimilis native streptokinase encoding DNA.
 XX
 KW Streptococcus equisimilis H46A; streptokinase; mutant; fibrinolytic;
 KW plasma clot; hydrolysis; haemolytic Streptococcus; plasminogen; plasmin;
 KW serine protease; fibrin; blood clot; thrombolytic;
 KW vascular thromboembolytic symptom; acute myocardial infarction;
 KW fibrinolysis; resistance; ds.
 XX
 OS Streptococcus dysgalactiae subsp. equisimilis.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1242
 FT /tag= a
 FT /transl_except= (pos:40..42,aa:Asn)
 FT /note= "no stop codon given"
 XX
 PN US5876999-A.
 XX
 PD 02-MAR-1999.
 XX
 PF 06-DEC-1995; 95US-00568393.
 XX
 PR 06-DEC-1995; 95US-00568393.
 XX
 PA (NASC-) NAT SCI COUNCIL.
 XX
 PI Wu H;
 XX
 DR WPI; 1999-189643/16.
 DR P-PSDB; AAW94664.
 XX
 PT Mutant streptokinase polypeptide - useful as plasmin-resistant
 PT thrombolytic agent.
 XX
 PS Claim 1; Col 7-10; 17pp; English.
 XX
 CC The present invention describes a mutant streptokinase (SK) polypeptide
 CC in which at least one of the amino acids in the Pro58-Lys59-Ser60-Lys61
 CC segment of the corresponding native SK is replaced by another amino acid.
 CC The present sequence encodes native SK. SK is a secretory protein of
 CC haemolytic Streptococcus able to activate human plasminogen (HPLg) to
 CC plasmin (HPLm), which is a serine protease able to catalyse the
 CC hydrolysis of fibrin in blood clots. The SK is useful as a thrombolytic
 CC agent in the treatment of vascular thromboembolytic symptoms such as
 CC acute myocardial infarction. Compared with wild-type SK, the X598 mutant
 CC is more resistant to degradation by human plasmin and is more effective
 CC both in acting as a fibrinolytic agent and in activating human plasminogen.
 CC (Updated on 17-OCT-2003 to standardise OS field)

SQ Sequence 1242 BP; 423 A; 268 C; 237 G; 314 T; 0 U; 0 Other;
 Query Match 69.1%; Score 1147; DB 2; Length 1242;
 Best Local Similarity 99.6%; Pred. No. 3.3e-314;
 Matches 1150; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 183 ATTCTCGACCTCGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCAATTTGTTGTTT 242
 Db 1 ATTCTCGACCTCGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCAATTTGTTGTTT 60
 QY 243 AGCGTTGCTGTACTGTGTGAGGGGACGAATCAAGACATTAGCTTTAAATTTTTTGAATC 302
 Db 61 AGCGTTGCTGTACTGTGTGAGGGGACGAATCAAGACATTAGCTTTAAATTTTTTGAATC 120
 QY 303 GATCTAACATCACGACCTGCTCATGGAGGAAAGACAGAGCAAGCTTAAGTCCAAAATCA 362
 Db 121 GATCTAACATCACGACCTGCTCATGGAGGAAAGACAGAGCAAGCTTAAGTCCAAAATCA 180
 QY 363 AAACCATTTGCTACTGTAGTGGCGGATGTCACTAAACTTGGAGAAAGCTGACTTACTA 422
 Db 181 AAACCATTTGCTACTGTAGTGGCGGATGTCACTAAACTTGGAGAAAGCTGACTTACTA 240
 QY 423 AAGGCTATTCAAGAACCAATTGATCGCTAAAGTCCACAGTAAACGACGACTCTTTGAGTGC 482
 Db 241 AAGGCTATTCAAGAACCAATTGATCGCTAAAGTCCACAGTAAACGACGACTCTTTGAGTGC 300
 QY 483 ATTGATTTTGCAGCGATGCAACCATTTACTGATCGAAACCGCAAGGCTACTTTGCTGAC 542
 Db 301 ATTGATTTTGCAGCGATGCAACCATTTACTGATCGAAACCGCAAGGCTACTTTGCTGAC 360
 QY 543 AAAGATGGTTCGGTAACTCTTCCGACCACTCTGCAAGAAATTTTGTCTAAGCGGACAT 602
 Db 361 AAAGATGGTTCGGTAACTCTTCCGACCACTCTGCAAGAAATTTTGTCTAAGCGGACAT 420
 QY 603 GTGCGCGTTAGACCATATAAGAAAAACCAATACAAAAACCAAGCGAAATCTGTTGANTG 662
 Db 421 GTGCGCGTTAGACCATATAAGAAAAACCAATACAAAAACCAAGCGAAATCTGTTGANTG 480
 QY 663 GAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGGCTCAAA 722
 Db 481 GAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGGCTCAAA 540
 QY 723 GATCTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTACTA 782
 Db 541 GATCTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTACTA 600
 QY 783 GCTCAAGCACAAGCATTTTAAACAAAAACCAAGCGGCTATACGATTTATGAACGTGAC 842
 Db 601 GCTCAAGCACAAGCATTTTAAACAAAAACCAAGCGGCTATACGATTTATGAACGTGAC 660
 QY 843 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 902
 Db 661 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 720
 QY 903 TTTTACTTACCGTGTAAATAATCGGGAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 962
 Db 721 TTTTACTTACCGTGTAAATAATCGGGAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 780
 QY 963 AATGAAGAAATAAAACACACCTGACCTGATCTCTGAGAAATATTACGCTCTTAAAAAAGGG 1022
 Db 781 AATGAAGAAATAAAACACACCTGACCTGATCTCTGAGAAATATTACGCTCTTAAAAAAGGG 840
 QY 1023 GAAAGCCGTATGATCCCTTTGATCGGAGTCACTTTGAAACTGTTTACCATCAAAATACGTT 1082
 Db 841 GAAAGCCGTATGATCCCTTTGATCGGAGTCACTTTGAAACTGTTTACCATCAAAATACGTT 900
 QY 1083 GATGTCGATACCAACGAAATGCTTAAAGTAGGAGAGCTCTTAAAGAGTAGGAGGTAAC 1142
 Db 901 GATGTCGATACCAACGAAATGCTTAAAGTAGGAGAGCTCTTAAAGAGTAGGAGGTAAC 960
 QY 1143 TTAGACTTTCAGAGATTTTATACGATCCCTCGTGAATAGGCTAAACTACTCTACAAATCTC 1202
 Db 961 TTAGACTTTCAGAGATTTTATACGATCCCTCGTGAATAGGCTAAACTACTCTACAAATCTC 1020

Db	3093	TCAGATGACACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAAGGAGAGAA	3152
Qy	1313	TGCTAGCTATCATTTAGCCGTTGGT	1337
Db	3153	TGCTAGCTATCATTTAGCCTATGAT	3177

Search completed: January 31, 2006, 18:17:50
Job time : 1005.11 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2006, 16:24:19 ; Search time 293.256 Seconds
(without alignments)
10068.109 Million cell updates/sec

Title: US-09-940-235-10
Perfect score: 1661
Sequence: 1 gcaaccgcagcctagcc.....gaataagctgtaccatctaa 1661

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/prodata/1/ina/1 COMB.seq:*
2: /cgn2_6/prodata/1/ina/5 COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B COMB.seq:*
5: /cgn2_6/prodata/1/ina/H COMB.seq:*
6: /cgn2_6/prodata/1/ina/PTUS COMB.seq:*
7: /cgn2_6/prodata/1/ina/PP COMB.seq:*
8: /cgn2_6/prodata/1/ina/RE COMB.seq:*
9: /cgn2_6/prodata/1/ina/backfilee1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1150	69.2	2385	3	US-09-211-542A-1
2	1148.6	69.2	1242	3	US-09-211-542A-5
3	1147	69.1	1242	2	US-08-568-393B-1
4	1143.8	68.9	1242	2	US-08-568-393B-2
5	1134.6	68.3	2566	2	US-08-488-940-19
6	1127.8	67.9	1245	2	US-07-703-778D-1
7	1127.6	67.9	1257	2	US-07-854-596B-25
8	1126.2	67.8	1512	2	US-07-854-596B-27
9	1124.6	67.7	1317	2	US-07-854-596B-18
10	1124.6	67.7	1335	2	US-07-854-596B-14
11	1124.6	67.7	1458	2	US-07-854-596B-42
12	1124.6	67.7	1467	2	US-07-854-596B-46
13	1124.6	67.7	2589	2	US-07-854-596B-34
14	1096.6	66.0	1122	3	US-09-374-038-13
15	1096.6	66.0	1122	3	US-09-658-179-13
16	1096.6	66.0	1158	3	US-09-374-038-14
17	1096.6	66.0	1158	3	US-09-658-179-14
18	1089.2	65.6	1209	3	US-09-374-038-11
19	1089.2	65.6	1209	3	US-09-658-179-11
20	1089.2	65.6	1245	3	US-09-374-038-12
21	1089.2	65.6	1245	3	US-09-658-179-12
22	1080.4	65.0	2253	2	US-07-854-596B-39
23	1078.8	64.9	1119	2	US-07-854-596B-30
24	976.2	58.8	1068	3	US-09-211-542A-11

25	973.2	58.6	2208	3	US-09-211-542A-3	Sequence 3, Appli
26	762.8	45.9	1262	9	5240845-3	Patent No. 5240845
27	757.8	45.6	1242	9	5240845-2	Patent No. 5240845
28	450	27.1	450	3	US-09-211-542A-13	Sequence 13, Appl
29	316.8	19.1	7679	3	US-09-220-132-38	Sequence 38, Appl
30	316.8	19.1	7680	3	US-09-023-655-1289	Sequence 1289, Ap
31	316.8	19.1	7680	6	PCT-US95-09819-6	Sequence 6, Appli
32	316.8	19.1	7705	2	US-08-259-569-16	Sequence 16, Appl
33	316.8	19.1	7705	2	US-08-826-885-16	Sequence 16, Appl
34	316.8	19.1	7705	9	5455158-2	Patent No. 5455158
35	316.8	19.1	7803	2	US-08-551-356-1	Sequence 1, Appli
36	316.8	19.1	7803	6	PCT-US93-12887-1	Sequence 1, Appli
37	310.4	18.7	8044	3	US-09-566-921-135	Sequence 135, App
38	178	10.7	4100	3	US-09-813-718-5	Sequence 5, Appli
39	178	10.7	4682	3	US-09-813-718-3	Sequence 3, Appli
40	178	10.7	4682	3	US-09-813-718-7	Sequence 7, Appli
41	178	10.7	4742	3	US-09-813-718-15	Sequence 15, Appl
42	178	10.7	4811	3	US-09-813-718-13	Sequence 13, Appl
43	178	10.7	4877	3	US-09-813-718-11	Sequence 11, Appl
44	178	10.7	5018	3	US-09-813-718-9	Sequence 9, Appli
45	178	10.7	5174	3	US-09-813-718-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-211-542A-1
; Sequence 1, Application US/09211542A
; Patent No. 6210667
; GENERAL INFORMATION:
; APPLICANT: Reed, Guy L.
; TITLE OF INVENTION: BACTERIAL FIBRIN-DEPENDENT PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROMBERG & SUNSTEIN, LLP
; STREET: 125 Summer Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 15-December-1998
; CLASSIFICATION: 1653
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/069,497
; FILING DATE: 15-December-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Attorney, Strimpel, Harriet M.
; REGISTRATION NUMBER: 37,008
; REFERENCE/DOCKET NUMBER: 1874/111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)443-9292
; TELEFAX: (617)443-0004
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2385 base pairs
; TYPE: nucleic acid
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2385
US-09-211-542A-1

Query Match 69.2%; Score 1150; DB 3; Length 2385;

Best Local Similarity 99.6%; Pred. No. 0;
Matches 1153; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 180 ATGATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCCCAATGGTT 239
DB 1141 AGGATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCCCAATGGTT 1200

QY 240 GTTAGCGTTGCTGCTAGTGTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGGAA 299
DB 1201 GTTAGCGTTGCTGCTAGTGTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGGAA 1260

QY 300 ATCGATCTAACATCAAGCACTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAGTCCAAAA 359
DB 1261 ATCGATCTAACATCAAGCACTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAGTCCAAAA 1320

QY 360 TCAAAACCAATTTGCTACTGATAGTGGCGGATGTCATAACTTGAGAAAGCTGACTTA 419
DB 1321 TCAAAACCAATTTGCTACTGATAGTGGCGGATGTCATAACTTGAGAAAGCTGACTTA 1380

QY 420 CTAAAGGCTATTCAAGCAATTTGATCGCTTAACGTCACAGTAACGACGACTACTTTGAG 479
DB 1381 CTAAAGGCTATTCAAGCAATTTGATCGCTTAACGTCACAGTAACGACGACTACTTTGAG 1440

QY 480 GTCAATTGATTTTCAAGCGATGCAACCAATTAATGATCGAAACGGCAAGGCTACTTTGCT 539
DB 1441 GTCAATTGATTTTCAAGCGATGCAACCAATTAATGATCGAAACGGCAAGGCTACTTTGCT 1500

QY 540 GACAAAGATGTTGCTGTAACCTTGGCGACCAACCTGTCACAGAAATTTTGGTAAAGCGGA 599
DB 1501 GACAAAGATGTTGCTGTAACCTTGGCGACCAACCTGTCACAGAAATTTTGGTAAAGCGGA 1560

QY 600 CATGTGCGGTTAGACCATATAAGAAACCAATCAAAACCAAGCGAATCTGTTGAT 659
DB 1561 CATGTGCGGTTAGACCATATAAGAAACCAATCAAAACCAAGCGAATCTGTTGAT 1620

QY 660 GTGGAATATACGTGACAGTTTACTCCCTTAAACCCCTGATGACGATTTTACAGCAGGCTC 719
DB 1621 GTGGAATATACGTGACAGTTTACTCCCTTAAACCCCTGATGACGATTTTACAGCAGGCTC 1680

QY 720 AAGATATACGATTTGAAACACATGATGATCGGTGACACCATCATCATCTCAAGAAATTA 779
DB 1681 AAGATATACGATTTGAAACACATGATGATCGGTGACACCATCATCATCTCAAGAAATTA 1740

QY 780 CTAGCTCAAGCAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTATGACGT 839
DB 1741 CTAGCTCAAGCAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTATGACGT 1800

QY 840 GACTCTCAATCGTCACTCATGCAATGACATTTTCCGTACGATTTTACCAATGGATCAA 899
DB 1801 GACTCTCAATCGTCACTCATGCAATGACATTTTCCGTACGATTTTACCAATGGATCAA 1860

QY 900 GAGTTTACTTACCGTTTAAAAATCGGAAACAAAGCTTTATAGGATCAATAAAAAATCTGGT 959
DB 1861 GAGTTTACTTACCGTTTAAAAATCGGAAACAAAGCTTTATAGGATCAATAAAAAATCTGGT 1920

QY 960 CTGAATGAAGAAATAAACAACCTGATCTCTGAGAAATATTACGTCCTTAAAAA 1019
DB 1921 CTGAATGAAGAAATAAACAACCTGATCTCTGAGAAATATTACGTCCTTAAAAA 1980

QY 1020 GGGGAAAGCCGATGATCCCTTTGATCGGAGTCACTTTGAAACTGTTCACCATCAATATC 1079
DB 1981 GGGGAAAGCCGATGATCCCTTTGATCGGAGTCACTTTGAAACTGTTCACCATCAATATC 2040

QY 1080 GTTGATGTCGATACCAACGAATTTGCTTAAAGTGAGCAGCTCTTAAACAGCTAGCGAAGCT 1139
DB 2041 GTTGATGTCGATACCAACGAATTTGCTTAAAGTGAGCAGCTCTTAAACAGCTAGCGAAGCT 2100

QY 1140 AACTTAGCTTACAGATTTATACGATTCCTCGTGATAAGGCTAAACTTACTCTCAACAAT 1199
DB 2101 AACTTAGCTTACAGATTTATACGATTCCTCGTGATAAGGCTAAACTTACTCTCAACAAT 2160

QY 1200 CTCGATGCTTTTGGTATTGACTATACCTTAACCTGGAAGTACGAGGATATCCAGT 1259

DB 2161 CTCGATGCTTTTGGTATTATGGACTATACCTTAACCTGAAAGTAGAGGATAATCAAGAT 2220

QY 1260 GACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAGGAGAGTGGTATG 1319
DB 2221 GACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAGGAGAGTGGTATG 2280

QY 1320 TATCATTTAGCCGGTGGT 1337
DB 2281 TATCATTTAGCCCTATGAT 2298

RESULT 2
US-09-211-542A-5
; Sequence 5, Application US/09211542A
; Patent No. 6210667
; GENERAL INFORMATION:
; APPLICANT: Reed, Guy L.
; TITLE OF INVENTION: BACTERIAL FIBRIN-DEPENDENT PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROMBERG & SUNSTEIN, LLP
; STREET: 125 Summer Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/09/211,542A
; APPLICATION NUMBER: 15-December-1998
; FILING DATE: 15-December-1998
; CLASSIFICATION: 1653
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/069,497
; FILING DATE: 15-December-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Attorney, Strimpel, Harriet M.
; REGISTRATION NUMBER: 37,008
; REFERENCE/DOCKET NUMBER: 1874/111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)443-9292
; TELEFAX: (617)443-0004
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1242 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1242
; US-09-211-542A-5

Query Match 69.2%; Score 1148.6; DB 3; Length 1242;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1151; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 183 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCCAATTTGGTT 242
DB 1 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCCAATTTGGTT 60

QY 243 AGCGTTCTGCTACTGTTGAGGGGACCAATCAACACATTAGTCTTAAATTTTGAATC 302
DB 61 AGCGTTCTGCTACTGTTGAGGGGACCAATCAACACATTAGTCTTAAATTTTGAATC 120

QY 303 GATCTAAACATCAACGACCTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAGTCCAAATCA 362
DB 121 GATCTAAACATCAACGACCTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAGTCCAAATCA 180

```
QY 363 AAACCAATTTGCTACTGATAGTGGCGCGATGTCAATAAACTTGAGAAAGTGACTTACTA 422
Db 181 AAACCAATTTGCTACTGATAGTGGCGCGATGTCAATAAACTTGAGAAAGTGACTTACTA 240
QY 423 AAGGCTATTCAAGAACCAATTTGATCGCTAAAGTCCACAGTACGACGACTACTTTGAGTGC 482
Db 241 AAGGCTATTCAAGAACCAATTTGATCGCTAAAGTCCACAGTACGACGACTACTTTGAGTGC 300
QY 483 ATTGATTTTGCAGCGATGCAACCAATTTACTGATCGAAACGGCAAGTCTACTTTGCTGAC 542
Db 301 ATTGATTTTGCAGCGATGCAACCAATTTACTGATCGAAACGGCAAGTCTACTTTGCTGAC 360
QY 543 AAAGATGGTTGGTAACTTTGCCGACCAACCTGTCCAAGAAATTTTGTCTAAGCGGACAT 602
Db 361 AAAGATGGTTGGTAACTTTGCCGACCAACCTGTCCAAGAAATTTTGTCTAAGCGGACAT 420
QY 603 GTGCGGTTAGACCATATAAAGAAAAACCAATACAAAACCAAGGAAATCTGTTGATGTG 662
Db 421 GTGCGGTTAGACCATATAAAGAAAAACCAATACAAAACCAAGGAAATCTGTTGATGTG 480
QY 663 GAATATACTCTACAGTTTACTCCCTTAAACCTGATGACGATTTTCAGACCAAGTCTCAA 722
Db 481 GAATATACTCTACAGTTTACTCCCTTAAACCTGATGACGATTTTCAGACCAAGTCTCAA 540
QY 723 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCAATCTCAAGAAATTA 782
Db 541 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCAATCTCAAGAAATTA 600
QY 783 GCTCAAGCAAAAGCAATTTTAAACAAAACCAACCCAGGCTATACGATTTATGAACGTGAC 842
Db 601 GCTCAAGCAAAAGCAATTTTAAACAAAACCAACCCAGGCTATACGATTTATGAACGTGAC 660
QY 843 TCCCTCAATCGTCACTCATGCAATGACATTTTCGTTACGATTTTACCAATGATCAAGAG 902
Db 661 TCCCTCAATCGTCACTCATGCAATGACATTTTCGTTACGATTTTACCAATGATCAAGAG 720
QY 903 TTTTACTTACCGTGTAAAAATCGGGAACAAAGCTTATAGGATCAATAAAAAATCTGGTCTG 962
Db 721 TTTTACTTACCGTGTAAAAATCGGGAACAAAGCTTATAGGATCAATAAAAAATCTGGTCTG 780
QY 963 AATGAAGAAATAAACACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAAGGG 1022
Db 781 AATGAAGAAATAAACACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAAGGG 840
QY 1023 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTACCAATCAATACGTT 1082
Db 841 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTACCAATCAATACGTT 900
QY 1083 GATGTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAACAGCTAGCGAACGTAAC 1142
Db 901 GATGTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAACAGCTAGCGAACGTAAC 960
QY 1143 TTAGACTTTCAGAGATTTATACGATCCCTCGTGATAAGGCTAACTCTACAAACATCTC 1202
Db 961 TTAGACTTTCAGAGATTTATACGATCCCTCGTGATAAGGCTAACTCTACAAACATCTC 1020
QY 1203 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAGAGTAGAGGATAATCAGCATGAC 1262
Db 1021 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAGAGTAGAGGATAATCAGCATGAC 1080
QY 1263 ACCAACGGTATCATACCGTTTATATGGGCAAGGACCCGAGAGAGAGATGCTAGCTAT 1322
Db 1081 ACCAACGGTATCATACCGTTTATATGGGCAAGGACCCGAGAGAGAGATGCTAGCTAT 1140
QY 1323 CATTTAGCCGGTGT 1337
Db 1141 CATTTAGCCCTATGAT 1155
```

RESULT 3

US-08-568-393B-1

; Sequence 1, Application US/08568393B

; Patent No. 5876999

```
; GENERAL INFORMATION:
; APPLICANT: Hua-Lin Wu
; APPLICANT: Guey-Yueh Shi
; TITLE OF INVENTION: Preparation of novel streptokinase
; TITLE OF INVENTION: mutants as improved thrombolytic agents
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jeing & Chang
; STREET: Two No. 5876999th Second Street, Suite 290
; CITY: San Jose
; STATE: California
; COUNTRY: USA
; ZIP: 95113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44MB
; MEDIUM TYPE: storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Wordperfect 6.1 on Window 3.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/568,393B
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Chi-Ping Chang
; REGISTRATION NUMBER: 37,798
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (408) 288-8585
; TELEFAX: (408) 288-8386
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1242 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE:
; HYPOTHETICAL: N
; ANTI-SENSE: N
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus equisimilis H46A
; INDIVIDUAL ISOLATE: Maik, H., Roe, B., and Ferretti, J. J.;
; INDIVIDUAL ISOLATE: "Nucleotide sequence of the streptokinase gene from Streptococcus equisimilis H46A" from Gene 34:357-362 (1985).
; CELL TYPE: Streptococcus equisimilis H46A
; US-08-568-393B-1
```

```
Query Match 69.1%; Score 1147; DB 2; Length 1242;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1150; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 183 ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAAACAGCCAAATTTGGTTCTT 242
Db 1 ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAAACAGCCAAATTTAGTTCTT 60
QY 243 AGCGTTGCTGATCTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 302
Db 61 AGCGTTGCTGATCTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 120
QY 303 GATCTAACATCAGCACTGCTCTGAGGAAAGCAGACGAGGCTTAAGTCCAAATCA 362
Db 121 GATCTAACATCAGCACTGCTCTGAGGAAAGCAGACGAGGCTTAAGTCCAAATCA 180
QY 363 AAACCAATTTGCTACTGATAGTGGCGCGATGTCAATAAACTTGAGAAAGTGACTTACTA 422
Db 181 AAACCAATTTGCTACTGATAGTGGCGCGATGTCAATAAACTTGAGAAAGTGACTTACTA 240
QY 423 AAGGCTATTCAAGAACCAATTTGATCGCTAAAGTCCACAGTACGACGACTACTTTGAGTGC 482
Db 241 AAGGCTATTCAAGAACCAATTTGATCGCTAAAGTCCACAGTACGACGACTACTTTGAGTGC 300
QY 483 ATTGATTTTGCAGCGATGCAACCAATTTACTGATCGAAACGGCAAGTCTACTTTGCTGAC 542
```

Db 301 ATTGATTTTGAAGCGATGCAACATTTACTGATCGAAACGCAAGGCTCTACTTTGCTGAC 360
 QY 543 AAAGATGGTTCGGTAACCTTGGCGACCCCAACCTGTCCAGAAATTTTGTCTAAGCGGACAT 602
 Db 361 AAAGATGGTTCGGTAACCTTGGCGACCCCAACCTGTCCAGAAATTTTGTCTAAGCGGACAT 420
 QY 603 GTGCGGTTAGACCATATATAAGAAAAAACAATACAAACCAAGCGAAATCTGTTGATGTG 662
 Db 421 GTGCGGTTAGACCATATATAAGAAAAAACAATACAAACCAAGCGAAATCTGTTGATGTG 480
 QY 663 GAATATAGTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGTCTCAAA 722
 Db 481 GAATATAGTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGTCTCAAA 540
 QY 723 GATACCTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATACTA 782
 Db 541 GATACCTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATACTA 600
 QY 783 GCTCAAGCAACAAGCAATTTTAAACAAAAACCAACCCAGGCTATACGATTTATGAACGTGAC 842
 Db 601 GCTCAAGCAACAAGCAATTTTAAACAAAAACCAACCCAGGCTATACGATTTATGAACGTGAC 660
 QY 843 TCTCAATCGTCACTCATGACATGACATTTTCGGTACGATTTTACCAATGGATCAAGAG 902
 Db 661 TCTCAATCGTCACTCATGACATGACATTTTCGGTACGATTTTACCAATGGATCAAGAG 720
 QY 903 TTTTACTACCGTGTAAATAATCGGAACAGCTTATAGGATCAATAAAAATCTGGTCTG 962
 Db 721 TTTTACTACCGTGTAAATAATCGGAACAGCTTATAGGATCAATAAAAATCTGGTCTG 780
 QY 963 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 1022
 Db 781 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 840
 QY 1023 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTTCAACATCAAAATACGTT 1082
 Db 841 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTTCAACATCAAAATACGTT 900
 QY 1083 GATGTCGATACCAAGAAATGCTTAAAGTGACAGCTCTTAAACAGTACGGAACGTAAAC 1142
 Db 901 GATGTCGATACCAAGAAATGCTTAAAGTGAGAGCTCTTAAACAGTACGGAACGTAAAC 960
 QY 1143 TTAGACTTTCAGAGATTTATACGATCTCTGAGTAAAGCTAAACTCTCTCAACAATCTC 1202
 Db 961 TTAGACTTTCAGAGATTTATACGATCTCTGAGTAAAGCTAAACTCTCTCAACAATCTC 1020
 QY 1203 GATGCTTTTGGTATTATGGAATATACCTTAACTGGAAGAGTAGAGGATATCAGATGAC 1262
 Db 1021 GATGCTTTTGGTATTATGGAATATACCTTAACTGGAAGAGTAGAGGATATCAGATGAC 1080
 QY 1263 ACCAACCGTATCATTAACCGTTTATATGGCAAGCGACCCGAGGAGAGAAATGCTAGCTAT 1322
 Db 1081 ACCAACCGTATCATTAACCGTTTATATGGCAAGCGACCCGAGGAGAGAAATGCTAGCTAT 1140
 QY 1323 CATTTAGCCGGTGGT 1337
 Db 1141 CATTTAGCCCTATGAT 1155

RESULT 4

US-08-568-393B-2
 ; Sequence 2, Application US/08568393B
 ; Patent No. 5876999
 ; GENERAL INFORMATION:
 ; APPLICANT: Hua-Lin Wu
 ; APPLICANT: Guey-Yueh Shi
 ; TITLE OF INVENTION: Preparation of novel streptokinase
 ; TITLE OF INVENTION: mutants as improved thrombolytic agents
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSER: Jeing & Chang
 ; STREET: Two No. 5876999th Second Street, Suite 290
 ; CITY: San Jose

STATE: California
 COUNTRY: USA
 ZIP: 95113
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.50 inch, 1.44MB
 MEDIUM TYPE: storage
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: WordPerfect 6.1 on Window 3.1
 CURRENT APPLICATION DATA: US/08/568,393B
 APPLICATION NUMBER: US/08/568,393B
 FILING DATE:
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Chi-Ping Chang
 REGISTRATION NUMBER: 37,798
 REFERENCE/DOCKET NUMBER:
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (408)288-8386
 TELEFAX: (408)288-8386
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1242 base pairs
 TYPE: Nucleic Acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE:
 HYPOTHETICAL: N
 ANTI-SENSE: N
 FEATURE:
 NAME/KEY: SK-K59B
 LOCATION: DNA sequence No. 5876999174 and 175 have been changed
 LOCATION: from AA to GG, and PROTEIN sequence No. 587699959 has been changed
 LOCATION: from lys to Glu.
 OTHER INFORMATION:
 US-08-568-393B-2

Query Match 68.9%; Score 1143.8; DB 2; Length 1242;

Best Local Similarity 99.4%; Pred. No. 0;
 Matches 1148; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 183 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCAATGCTGTTT 242
 Db 1 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCAATGCTGTTT 60
 QY 243 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 302
 Db 61 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 120
 QY 303 GATCTAATCATCAGACCTGCTCATGGAGGAAAGACAGACGAGCTTAAAGTCCAAATCA 362
 Db 121 GATCTAATCATCAGACCTGCTCATGGAGGAAAGACAGACGAGCTTAAAGTCCAAATCA 180
 QY 363 AAACCATTTGCTACTGATAGTGGCGCATGTCAATAAACTTGAGAAAGCTGACTTACTA 422
 Db 181 AAACCATTTGCTACTGATAGTGGCGCATGTCAATAAACTTGAGAAAGCTGACTTACTA 240
 QY 423 AAGGCTATTCAAGAACAAATTTGATCGCTAAACGTCACAGTAAACGACGACTACTTTGAGGTC 482
 Db 241 AAGGCTATTCAAGAACAAATTTGATCGCTAAACGTCACAGTAAACGACGACTACTTTGAGGTC 300
 QY 483 ATTGATTTTGCAGACGATGCAACCATTTACTGATCGAAACCGCAGGCTCTACTTTGCTGAC 542
 Db 301 ATTGATTTTGCAGACGATGCAACCATTTACTGATCGAAACCGCAGGCTCTACTTTGCTGAC 360
 QY 543 AAAGATGTTTCGGTAACTTTGCCGACCCAACTCTGCCAAGAAATTTTGTCTAAGCGGACAT 602
 Db 361 AAAGATGTTTCGGTAACTTTGCCGACCCAACTCTGCCAAGAAATTTTGTCTAAGCGGACAT 420
 QY 603 GTGCGGTTTAGACCATATAAAGAAAAAACAATACAAAAACCAAGCGAAATCTGTTGATGTG 662
 Db 421 GTGCGGTTTAGACCATATAAAGAAAAAACAATACAAAAACCAAGCGAAATCTGTTGATGTG 480

QY 663 GAATATCTGTACAGTTTACTCCCTTAAACCTGATGAGATTTTCAGACGAGTCTCAAA 722
Db 481 GAATATCTGTACAGTTTACTCCCTTAAACCTGATGAGATTTTCAGACGAGTCTCAAA 540
QY 723 GATACTAAGCTATTGAAACACTAGCTATCGGTGACCACTCATCTCAAGATTTACTA 782
Db 541 GATACTAAGCTATTGAAACACTAGCTATCGGTGACCACTCATCTCAAGATTTACTA 600
QY 783 GCTCAAGCACAAGCAATTTTAAACCAAAACCAACCCAGGCTATACGATTTATGAACGTGAC 842
Db 601 GCTCAAGCACAAGCAATTTTAAACCAAAACCAACCCAGGCTATACGATTTATGAACGTGAC 660
QY 843 TCCTCAATCGTCACTGATGACATGACATTTTCCGTAGCATTTTACCAATGATCAAGAG 902
Db 661 TCCTCAATCGTCACTGATGACATGACATTTTCCGTAGCATTTTACCAATGATCAAGAG 720
QY 903 TTTTACTTTACCGTGTAAATAATCGGGAACAGCTTATAGGATCAATATAAATAATCGGTCTG 962
Db 721 TTTTACTTTACCGTGTAAATAATCGGGAACAGCTTATAGGATCAATATAAATAATCGGTCTG 780
QY 963 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAGGG 1022
Db 781 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAGGG 840
QY 1023 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTTCAACATCAATACGTT 1082
Db 841 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTTCAACATCAATACGTT 900
QY 1083 GATGTCGATACCAACGAATTTGCTTAAAGTGAGGAGCTCTTAAACAGCTAGCGAAGCTAAC 1142
Db 901 GATGTCGATACCAACGAATTTGCTTAAAGTGAGGAGCTCTTAAACAGCTAGCGAAGCTAAC 960
QY 1143 TTAGACTTCAGAGATTTTATACGATCTCTGATGAGGCTAAACTACTCTACAACATCTC 1202
Db 961 TTAGACTTCAGAGATTTTATACGATCTCTGATGAGGCTAAACTACTCTACAACATCTC 1020
QY 1203 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAGAGTAGAGGATAATCAGCATGAC 1262
Db 1021 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAGAGTAGAGGATAATCAGCATGAC 1080
QY 1263 ACCAAGCTATCATACCGTTTATATGGGCAAGCGACCCGAGGAGAGAATGCTAGCTAT 1322
Db 1081 ACCAAGCTATCATACCGTTTATATGGGCAAGCGACCCGAGGAGAGAATGCTAGCTAT 1140
QY 1323 CATTTAGCCGGTGGT 1337
Db 1141 CATTTAGCCCTATGAT 1155

RESULT 5
US-08-488-940-19
; Sequence 19, Application US/08488940
; Patent No. 5854049
; GENERAL INFORMATION:
; APPLICANT: Reed, Guy L.
; TITLE OF INVENTION: PLASMIN-RESISTANT STREPTOKINASE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,940
; FILING DATE: 09-JUN-1995
; CLASSIFICATION: 514

; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 05433/009001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2566 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-488-940-19

Query Match 68.3%; Score 1134.6; DB 2; Length 2566;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1149; Conservative 0; Mismatches 4; Indels 2; Gaps 1;
QY 183 ATTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCAATGGTTGTT 242
Db 897 ATTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCAATGGTTGTT 956
QY 243 AGCGTTGCTGTTGAGGGGAGCAATCAAGACATTTAGTCTTAAATTTTTCGAATC 302
Db 957 AGCGTTGCTGTTGAGGGGAGCAATCAAGACATTTAGTCTTAAATTTTTCGAATC 1016
QY 303 GATCTAACATCAACGACCTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA 362
Db 1017 GATCTAACATCAACGACCTGCTCAT--AGGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA 1074
QY 363 AAACCATTTGCTATGATAGTGGCGCGATGTCAATAAACTTGAGAAAGCTGACTTACTA 422
Db 1075 AAACCATTTGCTATGATAGTGGCGCGATGTCAATAAACTTGAGAAAGCTGACTTACTA 1134
QY 423 AAGCTATTCAAGACAAATTCGTTAACTGATCGTAACTGCAAGCGCAAGGCTTCTTTGAGTGC 482
Db 1135 AAGCTATTCAAGACAAATTCGTTAACTGATCGTAACTGCAAGCGCAAGGCTTCTTTGAGTGC 1194
QY 483 ATTGATTTTGAAGCGATGCAACCAATTAATGATCGTAACTGCAAGCGCAAGGCTTCTTTGAGTGC 542
Db 1195 ATTGATTTTGAAGCGATGCAACCAATTAATGATCGTAACTGCAAGCGCAAGGCTTCTTTGAGTGC 1254
QY 543 AAAGATGGTTCGGTAACTTCCCGACCAACCTGTCGAAGAAATTTTTCGAAGCGGACAT 602
Db 1255 AAAGATGGTTCGGTAACTTCCCGACCAACCTGTCGAAGAAATTTTTCGAAGCGGACAT 1314
QY 603 GTGCGGTTAGACCATATAAAGAAAAACCAATATACAAACCAAGCGAAATCTGTTGATGTG 662
Db 1315 GTGCGGTTAGACCATATAAAGAAAAACCAATATACAAACCAAGCGAAATCTGTTGATGTG 1374
QY 663 GAATATCTGTACAGTTTACTCCCTTAAACCTGATGAGATTTTCAGACCAAGTCTCAAA 722
Db 1375 GAATATCTGTACAGTTTACTCCCTTAAACCTGATGAGATTTTCAGACCAAGTCTCAAA 1434
QY 723 GATACTAAGCTATTGAAACACTAGCTATCGGTGACCACTCATCTCAAGAAATTTACTA 782
Db 1435 GATACTAAGCTATTGAAACACTAGCTATCGGTGACCACTCATCTCAAGAAATTTACTA 1494
QY 783 GCTCAAGCACAAGCAATTTTAAACCAAAACCAACCCAGGCTATACGATTTATGAACGTGAC 842
Db 1495 GCTCAAGCACAAGCAATTTTAAACCAAAACCAACCCAGGCTATACGATTTATGAACGTGAC 1554
QY 843 TCCTCAATCGTCACTGATGACATGACATTTTCCGTAGCATTTTACCAATGATCAAGAG 902
Db 1555 TCCTCAATCGTCACTGATGACATGACATTTTCCGTAGCATTTTACCAATGATCAAGAG 1614
QY 903 TTTTACTTTACCGTGTAAATAATCGGGAACAGCTTATAGGATCAATATAAATAATCGGTCTG 962
Db 1615 TTTTACTTTACCGTGTAAATAATCGGGAACAGCTTATAGGATCAATATAAATAATCGGTCTG 1674

QY 963 AATGAAGAAATAAACAACACGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1022
 Db 1675 AATGAAGAAATAAACAACACGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1734
 QY 1023 GAAAACCGTATGATCCCTTTGATCGGAGTCACTTGAACCTGTTCAACATCAATACGTT 1082
 Db 1735 GAAAACCGTATGATCCCTTTGATCGGAGTCACTTGAACCTGTTCAACATCAATACGTT 1794
 QY 1083 GATGTCGATACCAACGAATTTGCTAAAAAGTGACGCTCTTAAACGCTAGCGAACGTAAC 1142
 Db 1795 GATGTCGATACCAACGAATTTGCTAAAAAGTGACGCTCTTAAACGCTAGCGAACGTAAC 1854
 QY 1143 TTAGACTTCAGAGATTTATACGATCTCTGATGAAGGCTAAACTACTCTACAACTCTC 1202
 Db 1855 TTAGACTTCAGAGATTTATACGATCTCTGATGAAGGCTAAACTACTCTACAACTCTC 1914
 QY 1203 GATGCTTTTGGTATTAAGACTATACCTTAACTGGAAGGCTAGAGGATTAATCAGATGAC 1262
 Db 1915 GATGCTTTTGGTATTAAGACTATACCTTAACTGGAAGGCTAGAGGATTAATCAGATGAC 1974
 QY 1263 ACCAACCGTATCATAACCGTTTATATGCGCAAGGACCCGCAAGGAGAGATGCTAGCTAT 1322
 Db 1975 ACCAACCGTATCATAACCGTTTATATGCGCAAGGACCCGCAAGGAGAGATGCTAGCTAT 2034
 QY 1323 CATTTAGCCGGTGGT 1337
 Db 2035 CATTTAGCCCTATGAT 2049

RESULT 6

US-07-703-778D-1
 ; Sequence 1, Application US/07703778D
 ; Patent No. 5296366
 ; GENERAL INFORMATION:
 ; APPLICANT: Garcia, M. P. E. et al
 ; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND EXPRESSION
 ; TITLE OF INVENTION: OF A GENE WHICH CODES FOR STREPTOKINASE, NUCLEOTIDE SEQUENCE
 ; TITLE OF INVENTION: OBTAINED, RECOMBINANT DNA AND TRANSFORMED MICROORGANISMS
 ; NUMBER OF SEQUENCES: 14
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSER: Stanger, Michaelson, Spivak and Wallace, Esq.
 ; STREET: Parkway 109 Office Center, 328 Newman Springs Road,
 ; STREET: P. O. Box 8489
 ; CITY: Red Bank
 ; STATE: New Jersey
 ; COUNTRY: USA
 ; ZIP: 07701
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3 1/2" 1.44Mb IBM compatible diskette
 ; COMPUTER: IBM PS/2 Model 80
 ; OPERATING SYSTEM: MS-DOS 5.0
 ; SOFTWARE: Microsoft Word for Windows 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/703,778D
 ; FILING DATE: 19910522
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Michaelson, Peter L.
 ; REGISTRATION NUMBER: 30090
 ; REFERENCE/DOCKET NUMBER: Centro-1
 ; TELEPHONE: (908)530-6671
 ; TELEFAX: (908)530-6584
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; TYPE: NUCLEOTIDE WITH CORRESPONDING PROTEIN
 ; LENGTH: 1245 base pairs
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: genomic DNA
 ; ORIGINAL SOURCE:
 ; ORGANISM: Streptococcus equisimilis from group C of Lanfield
 ; ORGANISM: definition

IMMEDIATE SOURCE: ATCC-9542 strain
 FEATURE: from 1 to 1245 bp mature peptide
 OTHER INFORMATION: Properties: Streptokinase gene
 OTHER INFORMATION: The gene product binds to human plasminogen
 OTHER INFORMATION: The gene product is an activator of human plasminogen
 US-07-703-778D-1
 Query Match 67.9%; Score 1127.8; DB 2; Length 1245;
 Best Local Similarity 98.5%; Pred. No. 0;
 Matches 1138; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
 QY 183 ATTGCTCGACCTGAGTGGCTCTAGACCGTCCATCTGTCAACACAGCAATTTGGTTGTT 242
 Db 1 ATTGCTGGACCTGAGTGGCTCTAGACCGTCCATCTGTCAACACAGCAATTTAGTTGTT 60
 QY 243 AGCGTTCTGGTACTGTTGAGGGGACGAATCAAGACATTAAGTTTAAATTTTGAATTC 302
 Db 61 AGCGTTCTGGTACTGTTGAGGGGACGAATCAAGACATTAAGTTTAAATTTTGAATTC 120
 QY 303 GATCTAATCATCAGACCTGCTCATGGAGGAAGACAGAGCAAGGCTTAAGTCCAAATCA 362
 Db 121 GACCTAAATCATCAGACCTGCTCATGGAGGAAGACAGAGCAAGGCTTAAGTCCAAATCA 180
 QY 363 AAACCATTTGCTACTGATAGTGGCGCGATGTACATAAACTTGAGAAAGCTGACTTACTA 422
 Db 181 AAACCATTTGCTACTGATAGTGGCGCGATGTACATAAACTTGAGAAAGCTGACTTACTA 240
 QY 423 AAGGCTATTCAAGAAACAATTAATGATCGCTTAAACGTCACAGTAAACGACGACTACTTTGAGTTC 482
 Db 241 AAGGCTATTCAAGAAACAATTAATGATCGCTTAAACGTCACAGTAAACGACGACTACTTTGAGTTC 300
 QY 483 ATTGATTTTGCAGCGATGCAACCATTAAGTCAAGGCTTAAAGTCCAAAGGCTTACTTTGCTGAC 542
 Db 301 ATTGATTTTGCAGCGATGCAACCATTAAGTCAAGGCTTAAAGTCCAAAGGCTTACTTTGCTGAC 360
 QY 543 AAAGATGTTTCGGTAAACCTTCCGACCCCAACCTGTCGAAGAAATTTTGTCTAAGCGGACAT 602
 Db 361 AAAGATGTTTCGGTAAACCTTCCGACCCCAACCTGTCGAAGAAATTTTGTCTAAGCGGACAT 420
 QY 603 GTGCGCTTTAGACCATATAAAGAAAAACCAATACAAAAACCAAGCGAAATCTGTTGATGTG 662
 Db 421 GTGCGCTTTAGACCATATAAAGAAAAACCAATACAAAAACCAAGCGAAATCTGTTGATGTG 480
 QY 663 GAATATCTGTACAGTTTACTTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTCAAA 722
 Db 481 GAATATCTGTACAGTTTACTTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTCAAA 540
 QY 723 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTA 782
 Db 541 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTA 600
 QY 783 GCTCAAGCACAAGCAATTTTAAAAACCAAAACCAAGGCTATACGATTTATGAAACGTCAC 842
 Db 601 GCTCAAGCACAAGCAATTTTAAAAACCAAAACCAAGGCTATACGATTTATGAAACGTCAC 660
 QY 843 TCCTCAATCGTCACTCATGACCAATGACATTTTCCGTAGAGTTTACCAATGATCAAGAG 902
 Db 661 TCCTCAATCGTCACTCATGACCAATGACATTTTCCGTAGAGTTTACCAATGATCAAGAG 720
 QY 903 TTTTACTTACCGTGTAAAAATCGGGAACAGCTTATAGGATCAATAAAAATCTGGTCTG 962
 Db 721 TTTTACTTACCGTGTAAAAATCGGGAACAGCTTATAGGATCAATAAAAATCTGGTCTG 780
 QY 963 AATGAAGAAATAAACAACACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1022
 Db 781 AATGAAGAAATAAACAACACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 840
 QY 1023 GAAAACCGTATGATCCCTTTGATCGGAGTCACTTGAACCTGTTCAACATCAATACGTT 1082
 Db 841 GAAAACCGTATGATCCCTTTGATCGGAGTCACTTGAACCTGTTCAACATCAATACGTT 900
 QY 1083 GATGTCGATACCAAGATTTGCTAAAAAGTGACGAGCTCTTAAACGCTAGCGAACGTAAC 1142

Db 901 GATGTCAACACCAACGAATTTGCTTAAAGGCGAGCAGCTCTTAAAGCTAGCGAACTAAC 960
QY 1143 TTAGACTTCAGAGATTTATACGATCCTCGTGATAGGCTAACTACTCTACACAACTCTC 1202
Db 961 TTAGACTTCAGAGATTTATACGATCCTCGTGATAGGCTAACTACTCTACACAACTCTC 1020
QY 1203 GATGCTTTTGGTATTATGGAATCTATACCTTAACTGGAAGAGTAGAGGATAATCAGATGAC 1262
Db 1021 GATGCTTTTGGTATTATGGAATCTATACCTTAACTGGAAGAGTAGAGGATAATCAGATGAC 1080
QY 1263 ACCAACCGTATCATACCGTTTATATGCGCAAGCGACCGGAGAGAGATGCTAGCTAT 1322
Db 1081 ACCAACCGTATCATACCGTTTATATGCGCAAGCGACCGGAGAGAGATGCTAGCTAT 1140
QY 1323 CATTTAGCCGGTGGT 1337
Db 1141 CATTTAGCCGTATGAT 1155

RESULT 7

US-07-854-596B-25
; Sequence 25, Application US/07854596B
; Patent No. 5434073
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith M
; APPLICANT: Hunter, Michael G
; APPLICANT: Czaplowski, Lloyd G
; TITLE OF INVENTION: Proteins and nucleic acids
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Dr. John J. McDonnell
; STREET: Ten South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/854,596B
; FILING DATE: 03-JUN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,337
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1257 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..1257
; OTHER INFORMATION: /note= "Methionyl-streptokinase"
; OTHER INFORMATION: fusion protein"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 4..1248
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 4..1248
US-07-854-596B-25

Query Match 67.9%; Score 1127.6; DB 2; Length 1257;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 1139; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
QY 180 ATGATTGCTGGACCTGAGTGGCTGTAGACCGTCACTGTCAACACAGCCAAATTTGTT 239
Db 4 ATGATTGCTGGACCTGAGTGGCTGTAGACCGTCACTGTCAACACAGCCAAATTTGTT 63
QY 240 GTTAGCGTTGCTGGTACTGTTGAGGGGACGATCAAGACATTAAGTCTTAAATTTTGA 299
Db 64 GTTAGCGTTGCTGGTACTGTTGAGGGGACGATCAAGACATTAAGTCTTAAATTTTGA 123
QY 300 ATCGATCTAAACATCAGACCTGCTCATGAGGAAAGACAGACAGCGCTTAAAGTCCAAA 359
Db 124 ATTAGACCTTAACATCAGACCTGCTCATGAGGAAAGACAGACAGCGCTTAAAGTCCAAA 183
QY 360 TCAAAACCATTTGCTACTGATAGTGGCGGATGTCACTATAAACTTTGAGAAAGCTGACTTA 419
Db 184 TCAAAACCATTTGCTACTGATAGTGGCGGATGTCACTATAAACTTTGAAAGCTGACTTA 243
QY 420 CTAAAGGCTATTCAAGAACAAATTTGATCGCTAACGTCACAGTAAACGACGACTCTTTGAG 479
Db 244 CTAAAGGCTATTCAAGAACAAATTTGATCGCTAACGTCACAGTAAACGACGACTCTTTGAG 303
QY 480 GTCATTGATTTTGCAGCGATGCAACCATTTACTGATCGAAACGCGAAGTCTACTTTGCT 539
Db 304 GTCATTGATTTTGCAGCGATGCAACCATTTACTGATCGAAACGCGAAGTCTACTTTGCT 363
QY 540 GACAAAGATGTTGCTTAACCTTTGCCGACCCCAACCTGTCCAAGAAATTTTGTAAAGCGA 599
Db 364 GACAAAGATGTTGCTTAACCTTTGCCGACCCCAACCTGTCCAAGAAATTTTGTAAAGCGA 423
QY 600 CATGTGCGGTTAGACCATATAAAGAAAAACCAATACAAATCAAGCGAAATCTGTTGAT 659
Db 424 CATGTGCGGTTAGACCATATAAAGAAAAACCAATACAAATCAAGCGAAATCTGTTGAT 483
QY 660 GTGGAATATCTGTACAGTTTACTCCCTTAAACCTGTGATGAGGATTCAGACCGAGTCTC 719
Db 484 GTGGAATATCTGTACAGTTTACTCCCTTAAACCTGTGATGAGGATTCAGACCGAGTCTC 543
QY 720 AAAGATCTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCACTCAAGAAATTA 779
Db 544 AAAGATCTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCACTCAAGAAATTA 603
QY 780 CTAGCTCAAGCAAAAGCAATTTTAAACAAAAACCAACCCATCCAGGCTATACGATTTATGAACGT 839
Db 604 CTAGCTCAAGCAAAAGCAATTTTAAACAAAAACCAACCCATCCAGGCTATACGATTTATGAACGT 663
QY 840 GACTCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGATCAAA 899
Db 664 GACTCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGATCAAA 723
QY 900 GAGTTTACTTACCGTGTAAATAATCGGAAACAGCTTATAGATCAATAAATAATCTCGT 959
Db 724 GAGTTTACTTACCGTGTAAATAATCGGAAACAGCTTATAGATCAATAAATAATCTCGT 783
QY 960 CTGAATGAAGAAATAAACAACACTGACTGTCTGAGAAATATTTAGCTCTTAAAAA 1019
Db 784 CTGAATGAAGAAATAAACAACACTGACTGTCTGAGAAATATTTAGCTCTTAAAAA 843
QY 1020 GGGGAAAGCGGTATGATCCCTTTGATCGAGTCACTGTAAGCTTTCCCATCAATAC 1079
Db 844 GGGGAAAGCGGTATGATCCCTTTGATCGAGTCACTTGAAGCTTTCCCATCAATAC 903
QY 1080 GTTGATGTCGATACCAACGATTTCTTAAAGTGGAGCTCTTAAAGCTAGCGAAAGT 1139
Db 904 GTTGATGTCGATACCAACGATTTCTTAAAGTGGAGCTCTTAAAGCTAGCGAAAGT 963
QY 1140 AACTTTAGACTTCAGAGATTTTATACGATCCTCGTATAGGCTTAAAGTCTTCAACCAAT 1199
Db 964 AACTTTAGACTTCAGAGATTTTATACGATCCTCGTATAGGCTTAAAGTCTTCAACCAAT 1023

QY 1315 CTAGCTATCATTTAGCCGGTGGT 1337
Db 1394 CTAGCTATCATTTAGCCCTATGAT 1416

RESULT 9

US-07-854-596B-18
; Sequence 18, Application US/07854596B
; Patent No. 5434073
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith M
; APPLICANT: Hunter, Michael G
; APPLICANT: Czaplewski, Lloyd G
; TITLE OF INVENTION: Proteins and nucleic acids
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. John J. McDonnell
; STREET: Ten South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/854,596B
; FILING DATE: 03-JUN-1992

CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,337
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 18:

SEQUENCE CHARACTERISTICS:
; LENGTH: 1317 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..1317
; OTHER INFORMATION: /notes="OmpAL fused to mature
; OTHER INFORMATION: streptokinase gene"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 4..1308
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 4..1308
US-07-854-596B-18

Query Match 67.7%; Score 1124.6; DB 2; Length 1317;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 1136; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 183 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCCAATTGGTTGTT 242
Db 67 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCCAATTAGTTGTT 126
QY 243 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAATC 302
Db 127 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAAT 186
QY 303 GATCTAACATCAGCACTGCTCATGGAGGAAAGACAGACGCAAGCTTAAAGTCCAAATCA 362

Db 187 GACCTAAACATCACGACCTGCTCATGGAGAAAGACAGAGCAAGGCTTAAAGTCCAAATCA 246
QY 363 AAACCATTTGCTACTGATAGTGGCGGATGTACATATAACTTTGAGAAAGCTGACTACTA 422
Db 247 AAACCATTTGCTACTGATAGTGGCGGATGTACATATAACTTTGAGAAAGCTGACTACTA 306
QY 423 AAGGCTATTCAAGAACAAATTGATCGCTAAACGTCACAGTAAACGACGACTACTCTTCAGGTC 482
Db 307 AAGGCTATTCAAGAACAAATTGATCGCTAAACGTCACAGTAAACGACGACTACTCTTCAGGTC 366
QY 483 ATTGATTTTGAAGCGATGCAACCAATTACTGATCGAAACGCGCAAGGTCTACTTTGCTGAC 542
Db 367 ATTGATTTTGAAGCGATGCAACCAATTACTGATCGAAACGCGCAAGGTCTACTTTGCTGAC 426
QY 543 AAAGATGCTTGGTAAACCTTCCCGACCCCAACCTGTGCGAAGATTTTTCGTAAGCGGACAT 602
Db 427 AAAGATGCTTGGTAAACCTTCCCGACCCCAACCTGTGCGAAGATTTTTCGTAAGCGGACAT 486
QY 603 GTGGCGTTAGACCATATAAAGAAAAACCAATACAAAAACCAAGCGAAATCTGTTCATGTG 662
Db 487 GTGGCGTTAGACCATATAAAGAAAAACCAATACAAAAACCAAGCGAAATCTGTTCATGTG 546
QY 663 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCGAGGTCTCAAA 722
Db 547 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCGAGGTCTCAAA 606
QY 723 GATCTAAGCTATTGAAACACACTAGCTATCGGTGACCATCATCATCTCAAGAAATTACTA 782
Db 607 GATCTAAGCTATTGAAACACACTAGCTATCGGTGACCATCATCATCTCAAGAAATTACTA 666
QY 783 GCTCAAGCACAAAGCATTTTAAACAAAAACCAACCCAGGCTATACGATTTATGAACGTGAC 842
Db 667 GCTCAAGCACAAAGCATTTTAAACAAAAACCAACCCAGGCTATACGATTTATGAACGTGAC 726
QY 843 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTAGCATTTTACCAATGGATCAAGAG 902
Db 727 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTAGCATTTTACCAATGGATCAAGAG 786
QY 903 TTTACTTACCGTGTAAATAATCGGGAAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 962
Db 787 TTTACTTACCGTGTAAATAATCGGGAAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 846
QY 963 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAGCTCTTAAAAAAGGG 1022
Db 847 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAGCTCTTAAAAAAGGG 906
QY 1023 GAAAAGCGTATGATCCCTTTTGATCGCAGTCACTTTGAAACTGTTCACCATCAAAATACGTT 1082
Db 907 GAAAAGCGTATGATCCCTTTTGATCGCAGTCACTTTGAAACTGTTCACCATCAAAATACGTT 966
QY 1083 GATGTCGATACCAACGAATTGCTAAAAAGTGAGCAGCTCTTTAAACAGCTAGCGAAACGTAAC 1142
Db 967 GATGTCGATACCAACGAATTGCTAAAAAGTGAGCAGCTCTTTAAACAGCTAGCGAAACGTAAC 1026
QY 1143 TTAGACTTCAGAGATTTTATACGATCTCTGATAGGCTAAACTTACTCTACAAATCTC 1202
Db 1027 TTAGACTTCAGAGATTTTATACGATCTCTGATAGGCTAAACTTACTCTACAAATCTC 1086
QY 1203 GATCCTTTTGGTATTATGGACTATACCTTTAATCGGAAAGTAGAGGATAANTCAGGATGAC 1262
Db 1087 GATCCTTTTGGTATTATGGACTATACCTTTAATCGGAAAGTAGAGGATAANTCAGGATGAC 1146
QY 1263 ACCAACCGTATCATAAACCGTTTATATGGGCAAGGACCCCGAAGGAGAGAAATGCTAGCTAT 1322
Db 1147 ACCAACCGTATCATAAACCGTTTATATGGGCAAGGACCCCGAAGGAGAGAAATGCTAGCTAT 1206
QY 1323 CATTTAGCCGGTGGT 1337
Db 1207 CATTTAGCCCTATGAT 1221

RESULT 10

US-07-854-596B-14
 ; Sequence 14, Application US/07854596B
 ; Patent No. 5434073
 ; GENERAL INFORMATION:
 ; APPLICANT: Dawson, Keith M
 ; APPLICANT: Hunter, Michael G
 ; APPLICANT: Czaplowski, Lloyd G
 ; TITLE OF INVENTION: Proteins and nucleic acids
 ; NUMBER OF SEQUENCES: 73
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Dr. John J. McDonnell
 ; STREET: Ten South Wacker Drive, Suite 3000
 ; CITY: Chicago
 ; STATE: IL
 ; COUNTRY: USA
 ; ZIP: 60606
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/854,596B
 ; FILING DATE: 03-JUN-1992
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: McDonnell, John J
 ; REGISTRATION NUMBER: 26,949
 ; REFERENCE/DOCKET NUMBER: 92,337
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 312-715-1000
 ; TELEFAX: 312-715-1234
 ; TELEX: 910-221-5317
 ; INFORMATION FOR SEQ ID NO: 14:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1335 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: 1..1335
 ; OTHER INFORMATION: /note= "Streptokinase gene from S."
 ; OTHER INFORMATION: equisimilis"
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 7..1326
 ; FEATURE:
 ; NAME/KEY: mat_peptide
 ; LOCATION: 7..1326
 ; US-07-854-596B-14

Query Match 67.7%; Score 1124.6; DB 2; Length 1335;
 Best Local Similarity 98.4%; Pred. No. 0;
 Matches 1136; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
 QY 183 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAAATGGTGT 242
 DB 85 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAAATGGTGT 144
 QY 243 AGCGTGTGGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 302
 DB 145 AGCGTGTGGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAAT 204
 QY 303 GATCTAAATCAGACCTGCTCATGGAGGAAGACAGAGCAAGCTTTAAGTCCAAATCA 362
 DB 205 GACCTAAATCAGACCTGCTCATGGAGGAAGACAGAGCAAGCTTTAAGTCCAAATCA 264
 QY 363 AAACCAATTTGCTAGTATGTCGCGGATGTCAATAAACTTGAGAAAGCTGACTTACTA 422
 DB 265 AAACCAATTTGCTAGTATGTCGCGGATGTCAATAAACTTGAGAAAGCTGACTTACTA 324

QY 423 AAGGCTATTCAAGAACAAATTTGATCGCTAAACGCTCCACAGTAACGAGACTACTTTGAGTTC 482
 DB 325 AAGGCTATTCAAGAACAAATTTGATCGCTAAACGCTCCACAGTAACGAGACTACTTTGAGTTC 384
 QY 483 ATTGATTTTGAACGGATGCAACCAATTTACTGATCGAAACGGCAAGGTCTACTTTGCTGAC 542
 DB 385 ATTGATTTTGAACGGATGCAACCAATTTACTGATCGAAACGGCAAGGTCTACTTTGCTGAC 444
 QY 543 AAAGATGTTGCGGTAAACCTTCCGACCCCACTGCTCCAAAGAAATTTTTCGTAAGCGGACAT 602
 DB 445 AAAGATGTTGCGGTAAACCTTCCGACCCCACTGCTCCAAAGAAATTTTTCGTAAGCGGACAT 504
 QY 603 GTGCGCTTAGACCATATAAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 662
 DB 505 GTGCGCTTAGACCATATAAAGAAAAACCAATACAAATCAAGCGAAATCTGTTGATGTG 564
 QY 663 GAATATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTTCAGACCAAGTCTCAAA 722
 DB 565 GAATATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTTCAGACCAAGTCTCAAA 624
 QY 723 GATCTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATTA 782
 DB 625 GATCTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATTA 684
 QY 783 GCTCAAGCACAAAGCATTTTAAACAAAAACCCACGAGCTATACGATTTTATGAACGTGAC 842
 DB 685 GCTCAAGCACAAAGCATTTTAAACAAAAACCCACGAGCTATACGATTTTATGAACGTGAC 744
 QY 843 TCCTCAATCGTCACTCATGACCAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 902
 DB 745 TCCTCAATCGTCACTCATGACCAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 804
 QY 903 TTTACTTACCGTGTAAAAATCGGGAAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 962
 DB 805 TTTACTTTACCATGTCAAAAAATCGGGAAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 864
 QY 963 AATGAAGAAATAAACCAACACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 1022
 DB 865 AATGAAGAAATAAACCAACACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 924
 QY 1023 GAAAGCCGTATGATCCCTTTTGATCGCAGTCACTTTGAAACTGTTCACCATCAAAATACGTT 1082
 DB 925 GAAAGCCGTATGATCCCTTTTGATCGCAGTCACTTTGAAACTGTTCACCATCAAAATACGTT 984
 QY 1083 GATGTCATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTTAAACAGCTAGCGAACGTAAC 1142
 DB 985 GATGTCATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTTAAACAGCTAGCGAACGTAAC 1044
 QY 1143 TTAGACTTCAGAGATTTTATACGATCCCTGATGATAAGGCTAAACTACTCTACAACAAATCTC 1202
 DB 1045 TTAGACTTCAGAGATTTTATACGATCCCTGATGATAAGGCTAAACTACTCTACAACAAATCTC 1104
 QY 1203 GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAGAGTAGAGGATAATCAGATGAC 1262
 DB 1105 GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAGAGTAGAGGATAATCAGATGAC 1164
 QY 1263 ACCAACCGTATCAACCGTTTATATGGGCAAGCGACCCGGAAGGAGAGAAATGCTAGCTAT 1322
 DB 1165 ACCAACCGTATCAACCGTTTATATGGGCAAGCGACCCGGAAGGAGAGAAATGCTAGCTAT 1224
 QY 1323 CATTTAGCCGGTGGT 1337
 DB 1225 CATTTAGCCCTATGAT 1239

RESULT 11
 US-07-854-596B-42
 ; Sequence 42, Application US/07854596B
 ; Patent No. 5434073
 ; GENERAL INFORMATION:
 ; APPLICANT: Dawson, Keith M
 ; APPLICANT: Hunter, Michael G
 ; APPLICANT: Czaplowski, Lloyd G

; TITLE OF INVENTION: Proteins and nucleic acids
 ; NUMBER OF SEQUENCES: 73
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: Dr. John J. McDonnell
 ; STREET: Ten South Wacker Drive, Suite 3000
 ; CITY: Chicago
 ; STATE: IL
 ; COUNTRY: USA
 ; ZIP: 60606
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/854,596B
 ; FILING DATE: 03-JUN-1992
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: McDonnell, John J
 ; REGISTRATION NUMBER: 26,949
 ; REFERENCE/DOCKET NUMBER: 92,337
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 312-715-1000
 ; TELEFAX: 312-715-1234
 ; TELEX: 910-221-5317
 ; INFORMATION FOR SEQ ID NO: 42:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1458 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: 1..1458
 ; OTHER INFORMATION: /note= "Hirudin-streptokinase
 ; OTHER INFORMATION: fusion linked by Factor Xa cleavable IEGR"
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 1..1449
 ; NAME/KEY: mat_peptide
 ; LOCATION: 1..1449
 ; US-07-854-596B-42

Query Match 67.7%; Score 1124.6; DB 2; Length 1458;
 Best Local Similarity 98.4%; Pred. No. 0;
 Matches 1136; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY	183	ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCCAATTGGTTGTT	242
DB	208	ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCCAATTGGTTGTT	267
QY	243	AGCGTTCGTCTGCTACTGTTGAGGGGCAACATCAAGACATTAGTCTTAATTTTGAATC	302
DB	268	AGCGTTCGTCTGCTACTGTTGAGGGGCAACATCAAGACATTAGTCTTAATTTTGAATC	327
QY	303	GATCTAACATCAGACCTGCTCATGGAGGAAAGACAGACGAGCGTCTTAAGTCCAAATCA	362
DB	328	GACCTAACATCAGACCTGCTCATGGAGGAAAGACAGACGAGCGTCTTAAGTCCAAATCA	387
QY	363	AAACCAATTGCTACTGATAGTGGCGCATGTCAATAAATTGAGAAGCTGACTTACTA	422
DB	388	AAACCAATTGCTACTGATAGTGGCGCATGTCAATAAATTGAGAAGCTGACTTACTA	447
QY	423	AAGCTATTCAAGAACCAATTGATCGCTAAAGTCCACAGTAACGACGACTTGTAGTGC	482
DB	448	AAGCTATTCAAGAACCAATTGATCGCTAAAGTCCACAGTAACGACGACTTGTAGTGC	507
QY	483	ATTGATTTTCAAGCCGATGCAACCATTTACTGATCGAAACGCAAGGCTCTCTTGTGAC	542
DB	508	ATTGATTTTCAAGCCGATGCAACCATTTACTGATCGAAACGCAAGGCTCTCTTGTGAC	567

QY	543	AAAGATGGTTTCGGTAAACCTTGGCCGACCCAACTGCTCAAGAAATTTTGTGAAGCGGACAT	602
DB	568	AAAGATGGTTTCGGTAAACCTTGGCCGACCCAACTGCTCAAGAAATTTTGTGAAGCGGACAT	627
QY	603	GTGCGGTTAGACCATATAAAGAAAAACAATACAAAACCAAGCGAAATCTCTTGTATGTG	662
DB	628	GTGCGGTTAGACCATATAAAGAAAAACAATACAAAACCAAGCGAAATCTCTTGTATGTG	687
QY	663	GAATATCTGTACAGTTTACTCCCTTAACCCCTGATGACGATTCAGACCAGGTCTCAAA	722
DB	688	GAATATCTGTACAGTTTACTCCCTTAACCCCTGATGACGATTCAGACCAGGTCTCTCAA	747
QY	723	GATACTAAGCTATTGAAACACCTAGCTATCGGTGACACCATCACAATCTCAAGAAATCTA	782
DB	748	GATACTAAGCTATTGAAACACCTAGCTATCGGTGACACCATCACAATCTCAAGAAATCTA	807
QY	783	GCTCAAGCACAAAGCAATTTTAAACAAAACCAACCCAGGCTATACGATTTATGAAACGTAC	842
DB	808	GCTCAAGCACAAAGCAATTTTAAACAAAACCAACCCAGGCTATACGATTTATGAAACGTAC	867
QY	843	TCCTCAATCGTCATCATGACAATGACATTTTCCGTACGATTTTACCAATGATCAAGAG	902
DB	868	TCCTCAATCGTCATCATGACAATGACATTTTCCGTACGATTTTACCAATGATCAAGAG	927
QY	903	TTTACTTACCGTGTAAATAATCGGCAACAAGCTTATAGATCAATAAATAATCTCGTCTG	962
DB	928	TTTACTTACCGTGTAAATAATCGGCAACAAGCTTATAGATCAATAAATAATCTCGTCTG	987
QY	963	AATGAAGAAATAAACAACAACACTGACCTGATCTCTGAGAAATATTAGCTCTTAAAAAAGG	1022
DB	988	AATGAAGAAATAAACAACAACACTGACCTGATCTCTGAGAAATATTAGCTCTTAAAAAAGG	1047
QY	1023	GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTTGAAACCTGTTCCACATCAATACGTT	1082
DB	1048	GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTTGAAACCTGTTCCACATCAATACGTT	1107
QY	1083	GATGTCGATACCAAGCAATTTGCTTAAAGAGTGAAGCTCTTAAAGAGTGAAGCAATTAAC	1142
DB	1108	GATGTCGATACCAAGCAATTTGCTTAAAGAGTGAAGCTCTTAAAGAGTGAAGCAATTAAC	1167
QY	1143	TTAGACTTTCAGAGATTTATACGATCTCTGATAGGCTTAAACTACTCTACAAACATCTC	1202
DB	1168	TTAGACTTTCAGAGATTTATACGATCTCTGATAGGCTTAAACTACTCTACAAACATCTC	1227
QY	1203	GATGCTTTTGGTATTATGAGTACTATACCTTAACTGGAAAGTAGAGATAATCAGCATGAC	1262
DB	1228	GATGCTTTTGGTATTATGAGTACTATACCTTAACTGGAAAGTAGAGATAATCAGCATGAC	1287
QY	1263	ACCAACCGTATCATTAACCGTTTATATGGGCAAGGACCCGGAAGGAGAGAAATGCTAGCTAT	1322
DB	1288	ACCAACCGTATCATTAACCGTTTATATGGGCAAGGACCCGGAAGGAGAGAAATGCTAGCTAT	1347
QY	1323	CATTAGCCGGTGGT 1337	
DB	1348	CATTAGCCCTATGAT 1362	

RESULT 12
 US-07-854-596B-46
 ; Sequence 46, Application US/07854596B
 ; Patent No. 5434073
 ; GENERAL INFORMATION:
 ; APPLICANT: Dawson, Keith M
 ; APPLICANT: Hunter, Michael G
 ; APPLICANT: Czaplewski, Lloyd G
 ; TITLE OF INVENTION: Proteins and nucleic acids
 ; NUMBER OF SEQUENCES: 73
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Dr. John J. McDonnell
 ; STREET: Ten South Wacker Drive, Suite 3000
 ; CITY: Chicago
 ; STATE: IL

```

; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/854,596B
; FILING DATE: 03-JUN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,337
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1467 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..1467
; OTHER INFORMATION: /note= "Streptokinase-hirudin
; OTHER INFORMATION: fusion linked by Factor Xa-
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1449
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..1449
; US-07-854-596B-46

```

Query Match	67.7%	Score 1124.6	DB 2	Length 1467
Best Local Similarity	98.4%	Pred. No. 0		
Matches 1136	Conservative 0	Mismatches 19	Indels 0	Gaps 0
QY	183	ATTGCTCGACCTGAGTCGCTGCTAGACCGTCCATCTGTCAACAACAGCAATTCGTTGTT	242	
DB	1	ATTGCTCGACCTGAGTCGCTGCTAGACCGTCCATCTGTCAACAACAGCAATTCGTTGTT	60	
QY	243	AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTTTGAATC	302	
DB	61	AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTTTGAATC	120	
QY	303	GATCTAACATCAACGACCTGCTCATGAGGAGAAAGACAGACGCAAGGCTTAAAGTCCAAATCA	362	
DB	121	GACCTAACATCAACGACCTGCTCATGAGGAGAAAGACAGACGCAAGGCTTAAAGTCCAAATCA	180	
QY	363	AAACCAATTTGCTACTGATAGTGGCGCGATGTGCATAAACTTTGAGAAGCTGACTTACTA	422	
DB	181	AAACCAATTTGCTACTGATAGTGGCGCGATGTGCATAAACTTTGAGAAGCTGACTTACTA	240	
QY	423	AAGGCTATTCAAGAACCAATTGATCGCTAAACGTCCACAGTAAACGACGACTACTTTGAGGTC	482	
DB	241	AAGGCTATTCAAGAACCAATTGATCGCTAAACGTCCACAGTAAACGACGACTACTTTGAGGTC	300	
QY	483	ATTGATTTTGCAGCGATGCAACCAATTAATCTGATCGAAACGGCAGGCTCTACTTTGCTGAC	542	
DB	301	ATTGATTTTGCAGCGATGCAACCAATTAATCTGATCGAAACGGCAGGCTCTACTTTGCTGAC	360	
QY	543	AAAGATGGTTCGGTAAACCTTGCCGACCCAACTGTGCCAAGAAATTTTTTGTCTAAGCGGACAT	602	
DB	361	AAAGATGGTTCGGTAAACCTTGCCGACCCAACTGTGCCAAGAAATTTTTTGTCTAAGCGGACAT	420	
QY	603	GTGCGCGTTAGACCATATAAAGAAAAACCATAAAAAACCAAGCGAAATCTGTTGATGTG	662	

```

1  RESULT 13
2  US-07-854-596B-34
3  ; Sequence 34, Application US/07854596B
4  ; Patent No. 5434073
5  ; GENERAL INFORMATION:
6  ; APPLICANT: Dawson, Keith M
7  ; APPLICANT: Hunter, Michael G
8  ; APPLICANT: Czaplowski, Lloyd G
9  ; TITLE OF INVENTION: Proteins and nucleic acids
10 ; NUMBER OF SEQUENCES: 73
11 ; CORRESPONDENCE ADDRESS:
12 ; ADDRESSEE: Dr. John J. McDonnell
13 ; STREET: Ten South Wacker Drive, Suite 3000
14 ; CITY: Chicago
15 ; STATE: IL
16 ; COUNTRY: USA
17 ; ZIP: 60606
18 ; COMPUTER READABLE FORM:
19 ; MEDIUM TYPE: Floppy disk
20 ; COMPUTER: IBM PC compatible
21 ; OPERATING SYSTEM: PC-DOS/MS-DOS
22 ; SOFTWARE: PatentIn Release #1.0, Version #1.0

```

; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/854,596B
 ; FILING DATE: 03-JUN-1992
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: McDonnell, John J
 ; REGISTRATION NUMBER: 26,949
 ; REFERENCE/DOCKET NUMBER: 92,337
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 312-715-1000
 ; TELEFAX: 312-715-1234
 ; TELEX: 910-221-5317
 ; INFORMATION FOR SEQ ID NO: 34:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2589 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: 1..2589
 ; OTHER INFORMATION: /note=
 ; OTHER INFORMATION: "OmpAL-Streptokinase-streptokinase fusion linked
 ; OTHER INFORMATION: by thrombin-cleavable VELQGVPRG"
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 4..2580
 ; FEATURE:
 ; NAME/KEY: mat_peptide
 ; LOCATION: 4..2580
 ; US-07-854-596B-34

Query Match i 67.7%; Score 1124.6; DB 2; Length 2589;
 Best Local Similarity 98.4%; Pred. No. 0;
 Matches 1136; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 183 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCCAATTGGTGT 242
 DB 67 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCCAATTGGTGT 126
 QY 243 AGCTGTGCTGTTACTGTTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTGGAAATC 302
 DB 127 AGCTGTGCTGTTACTGTTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTGGAAAT 186
 QY 303 GATCTAATCATCAGCACTGCTCATGAGGAGGAGCAAGCAAGCTTTAAGTCCAAATCA 362
 DB 187 GACCTAATCATCAGCACTGCTCATGAGGAGGAGCAAGCAAGCTTTAAGTCCAAATCA 246
 QY 363 AAAACCAATTTGCTACTGATAGTGGCGGATGTCATATAAACTTGAGAAAGCTGACTTACTA 422
 DB 247 AAAACCAATTTGCTACTGATAGTGGCGGATGTCATATAAACTTGAGAAAGCTGACTTACTA 306
 QY 423 AAGGCTATTCAAGAACAAATTTGATCGCTTAACGTCACAGTACGACGACTACTTTGAGGTC 482
 DB 307 AAGGCTATTCAAGAACAAATTTGATCGCTTAACGTCACAGTACGACGACTACTTTGAGGTC 366
 QY 483 ATTGATTTTGAAGCGATGCAACCATTAATCTGATCGAAACGGCAAGGCTACTTTGCTGAC 542
 DB 367 ATTGATTTTGAAGCGATGCAACCATTAATCTGATCGAAACGGCAAGGCTACTTTGCTGAC 426
 QY 543 AAGATGTTTCGGTAACTTCCGCGCCCACTGTCAGCAATTTTTCGCTAAGCGGACAT 602
 DB 427 AAGATGTTTCGGTAACTTCCGCGCCCACTGTCAGCAATTTTTCGCTAAGCGGACAT 486
 QY 603 GTGCGGTTAGACCATATAAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 662
 DB 487 GTGCGGTTAGACCATATAAAGAAAAACCAATACAAATCAAGCGAAATCTGTTGATGTG 546
 QY 663 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAA 722
 DB 547 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAA 606

QY 723 GATCTAAGCTATTGAAAAACACTAGCTATCGGTGACACATCACAATCTCAAGAAATTACTA 782
 DB 607 GATCTAAGCTATTGAAAAACACTAGCTATCGGTGACACATCACAATCTCAAGAAATTACTA 666
 QY 783 GCTCAAGCACAAAGCATTTTAAACAAAAACCAACCCAGGCTATACGATTTATGAACGTCAC 842
 DB 667 GCTCAAGCACAAAGCATTTTAAACAAAAACCAACCCAGGCTATACGATTTATGAACGTCAC 726
 QY 843 TCCTCAATCGTCACCTCATGACAATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 902
 DB 727 TCCTCAATCGTCACCTCATGACAATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 786
 QY 903 TTTACTTACCCTGTTTAAAAATCGGAAACAAAGCTTATAGGATCAATPAAAAAATCTCGTCTG 962
 DB 787 TTTACTTACCCTGTTTAAAAATCGGAAACAAAGCTTATAGGATCAATPAAAAAATCTCGTCTG 846
 QY 963 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAGCTCTTAAAAAAGGG 1022
 DB 847 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAGCTCTTAAAAAAGGG 906
 QY 1023 GAAAAAGCGGTATGATCCCTTTTATGCGCAGTCACTTTGAAACTGTTTCCACCATCAAAATACGTT 1082
 DB 907 GAAAAAGCGGTATGATCCCTTTTATGCGCAGTCACTTTGAAACTGTTTCCACCATCAAAATACGTT 966
 QY 1083 GATGTCGATACCAACGAATTGCTTAAAAAGTGAAGCAGCTCTTAAACAGCTAGCGAAACGTAAC 1142
 DB 967 GATGTCACACCAACGAATTGCTTAAAAAGCAGCAGCTCTTAAACAGCTAGCGAAACGTAAC 1026
 QY 1143 TTAGACTTCAGAGATTTTATAGGATCCCTGCTGATAGGCTTAACTCTTACAAATCTC 1202
 DB 1027 TTAGACTTCAGAGATTTTATAGGATCCCTGCTGATAGGCTTAACTCTTACAAATCTC 1086
 QY 1203 GATGCTTTTGGTATTATGAGTACTTAACTTAACTGGAAGAGTAGAGGATAATCAGATGAC 1262
 DB 1087 GATGCTTTTGGTATTATGAGTACTTAACTTAACTGGAAGAGTAGAGGATAATCAGATGAC 1146
 QY 1263 ACCAACCGTATCATAACCGTTTATATGGGCAAGGACCCGAGGAGAGAGATGCTAGCTAT 1322
 DB 1147 ACCAACCGTATCATAACCGTTTATATGGGCAAGGACCCGAGGAGAGAGATGCTAGCTAT 1206
 QY 1323 CATTTAGCCGCTGTT 1337
 DB 1207 CATTTAGCCCTATGAT 1221

RESULT 14
 US-09-374-038-13
 ; Sequence 13, Application US/09374038
 ; Patent No. 6309873
 ; GENERAL INFORMATION:
 ; APPLICANT: Madrazo, Isis Del Carmen Torrens
 ; APPLICANT: Garcia, Jose De Jesus De La Fuente
 ; APPLICANT: Ojalvo, Ariana Garcia
 ; APPLICANT: Menendez, Alina Seraleña
 ; APPLICANT: Escalona, Elder Pupo
 ; APPLICANT: Masso, Julio Raul Fernandez
 ; APPLICANT: Griego, Martha De Jesus Gonzalez
 ; TITLE OF INVENTION: STREPTOKINASE MUTANTS
 ; FILE REFERENCE: Sequence Listings 1-14 re: 976-5
 ; Patent No. 6309873
 ; CURRENT APPLICATION NUMBER: US/09/374,038
 ; CURRENT FILING DATE: 1999-08-13
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 13
 ; LENGTH: 1122
 ; TYPE: DNA
 ; ORGANISM: Streptococcus equisimilis
 ; US-09-374-038-13

Query Match 66.0%; Score 1096.6; DB 3; Length 1122;
 Best Local Similarity 98.7%; Pred. No. 5e-308;
 Matches 1105; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 180 ATGATGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCCCAATTGGTT 239
 Db 1 ATGATGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCCCAATTAGTT 60
 Qy 240 GTTAGCGTTGCTGCTAGTCTGTTGAGGGGAGCAATCAAGACATTAAGTCTTAAATTTTGGAA 299
 Db 61 GTTAGCGTTGCTGCTAGTCTGTTGAGGGGAGCAATCAAGACATTAAGTCTTAAATTTTGGAA 120
 Qy 300 ATCGATCTAACATCAGACCTGCTCATGGAGGAAGACAGACGAGCAAGCTTAAAGTCCAAA 359
 Db 121 ATGACCTTAACATCAGACCTGCTCATGGAGGAAGACAGACGAGCAAGCTTAAAGTCCAAA 180
 Qy 360 TCAAAACCAATTTGCTACTGATAGTGGCGGATGTCATATAAATCTGAGAAAGCTGACTTA 419
 Db 181 TCAAAACCAATTTGCTACTGATAGTGGCGGATGTCATATAAATCTGAGAAAGCTGACTTA 240
 Qy 420 TAAAGGCTTATCAAGAACAAATGATCGCTAACGCTCCACAGTAACGACGACTACTTTGAG 479
 Db 241 TAAAGGCTTATCAAGAACAAATGATCGCTAACGCTCCACAGTAACGACGACTACTTTGAG 300
 Qy 480 GTCAATGATTTGCAAGCGATGCAACCATTAATCTGATCGAAACGGCAAGTCTACTTTGCT 539
 Db 301 GTCAATGATTTGCAAGCGATGCAACCATTAATCTGATCGAAACGGCAAGTCTACTTTGCT 360
 Qy 540 GACAAAGATGTTGCGTAACTTGGCGACCTTGCACCACTGTCACGAATTTTGTCTAAAGCGGA 599
 Db 361 GACAAAGATGTTGCGTAACTTGGCGACCTTGCACCACTGTCACGAATTTTGTCTAAAGCGGA 420
 Qy 600 CATGTGCGGTTAGACCATATAAAGAAACCAATAACAAACCAAGCGAAATCTGTTGAT 659
 Db 421 CATGTGCGGTTAGACCATATAAAGAAACCAATAACAAACCAAGCGAAATCTGTTGAT 480
 Qy 660 GTGGAATATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTACAGACAGGCTC 719
 Db 481 GTGGAATATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTACAGACAGGCTC 540
 Qy 720 AAGATATCTAAGCTATTGAAACACTAGTATCGGTGACCACTCATCATCTCAAGAAATTA 779
 Db 541 AAGATATCTAAGCTATTGAAACACTAGTATCGGTGACCACTCATCATCTCAAGAAATTA 600
 Qy 780 CTAGCTCAAGCAACAAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTATGAACGT 839
 Db 601 CTAGCTCAAGCAACAAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTATGAACGT 660
 Qy 840 GACTCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGGATCAA 899
 Db 661 GACTCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGGATCAA 720
 Qy 900 GAGTTTACTTACCGTTTAAATCGGGAACAGCTTATAGGATCAATATAAATCTGGT 959
 Db 721 GAGTTTACTTACCGTTTAAATCGGGAACAGCTTATAGGATCAATATAAATCTGGT 780
 Qy 960 CTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGCTCTTAAAAAA 1019
 Db 781 CTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGCTCTTAAAAAA 840
 Qy 1020 GGGGAAAAGCCGTATGATCCCTTTGATCGAGTCACTTGAAACTGTTTCCACCATCAAAATAC 1079
 Db 841 GGGGAAAAGCCGTATGATCCCTTTGATCGAGTCACTTGAAACTGTTTCCACCATCAAAATAC 900
 Qy 1080 GTTGATGTCGATACCAACGAATGCTTAAAGTGACAGCTCTTAAAGTGACAGTACGAAACGT 1139
 Db 901 GTTGATGTCGATACCAACGAATGCTTAAAGTGACAGCTCTTAAAGTGACAGTACGAAACGT 960
 Qy 1140 AACTTAGACTTCAGAGATTTATAGGATCTCTGCTGATAAGCTAAACTACTCTACAAAT 1199
 Db 961 AACTTAGACTTCAGAGATTTATAGGATCTCTGCTGATAAGCTAAACTACTCTACAAAT 1020
 Qy 1200 CTCGATGCTTTTGGTATTAAGACTATACCTTAACCTGGAAGAGTAGAGGATATACGAT 1259
 Db 1021 CTCGATGCTTTTGGTATTAAGACTATACCTTAACCTGGAAGAGTAGAGGATATACGAT 1080

Qy 1260 GACACCAACCGTATCATACCGTTTATATGGCAAGCGA 1298
 Db 1081 GACACCAACCGTATCATACCGTTTATATGGCAAGCGA 1119
 RESULT 15
 US-09-658-179-13
 ; Sequence 13, Application US/09658179
 ; Patent No. 6413759
 ; GENERAL INFORMATION:
 ; APPLICANT: Madrazo, Isis Del Carmen Torrens
 ; APPLICANT: Garcia, Jose De Jesus De La Fuente
 ; APPLICANT: Ojalvo, Ariana Garcia
 ; APPLICANT: Menendez, Alina Seralena
 ; APPLICANT: Escalona, Elder Pupo
 ; APPLICANT: Masso, Julio Raul Fernandez
 ; APPLICANT: Griego, Martha De Jesus Gonzalez
 ; TITLE OF INVENTION: STREPTOKINASE MUTANTS
 ; FILE REFERENCE: Sequence Listings 1-14 re: 976-5
 ; Patent No. 6413759
 ; CURRENT APPLICATION NUMBER: US/09/658,179
 ; CURRENT FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 13
 ; LENGTH: 1122
 ; TYPE: DNA
 ; ORGANISM: Streptococcus equisimilis
 US-09-658-179-13

Query Match 66.0%; Score 1096.6; DB 3; Length 1122;
 Best Local Similarity 98.7%; Pred. No. 5e-308;
 Matches 1105; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
 Qy 180 ATGATGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCCCAATTGGTT 239
 Db 1 ATGATGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCCCAATTAGTT 60
 Qy 240 GTTAGCGTTGCTGCTAGTCTGTTGAGGGGAGCAATCAAGACATTAAGTCTTAAATTTTGGAA 299
 Db 61 GTTAGCGTTGCTGCTAGTCTGTTGAGGGGAGCAATCAAGACATTAAGTCTTAAATTTTGGAA 120
 Qy 300 ATCGATCTAACATCAGACCTGCTCATGGAGGAAGACAGACGAGCAAGCTTAAAGTCCAAA 359
 Db 121 ATGACCTTAACATCAGACCTGCTCATGGAGGAAGACAGACGAGCAAGCTTAAAGTCCAAA 180
 Qy 360 TCAAAACCAATTTGCTACTGATAGTGGCGGATGTCATATAAATCTGAGAAAGCTGACTTA 419
 Db 181 TCAAAACCAATTTGCTACTGATAGTGGCGGATGTCATATAAATCTGAGAAAGCTGACTTA 240
 Qy 420 CTAAAGGCTTATCAAGAACAAATGATCGCTAACGCTCCACAGTAACGACGACTACTTTGAG 479
 Db 241 CTAAAGGCTTATCAAGAACAAATGATCGCTAACGCTCCACAGTAACGACGACTACTTTGAG 300
 Qy 480 GTCAATGATTTGCAAGCGATGCAACCATTAATCTGATCGAAACGGCAAGTCTACTTTGCT 539
 Db 301 GTCAATGATTTGCAAGCGATGCAACCATTAATCTGATCGAAACGGCAAGTCTACTTTGCT 360
 Qy 540 GACAAAGATGTTGCGTAACTTGGCGACCTTGCACCACTGTCACGAATTTTGTCTAAAGCGGA 599
 Db 361 GACAAAGATGTTGCGTAACTTGGCGACCTTGCACCACTGTCACGAATTTTGTCTAAAGCGGA 420
 Qy 600 CATGTGCGGTTAGACCATATAAAGAAACCAATAACAAACCAAGCGAAATCTGTTGAT 659
 Db 421 CATGTGCGGTTAGACCATATAAAGAAACCAATAACAAACCAAGCGAAATCTGTTGAT 480
 Qy 660 GTGGAATATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTACAGACAGGCTC 719
 Db 481 GTGGAATATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTACAGACAGGCTC 540
 Qy 720 AAGATATCTAAGCTATTGAAACACTAGTATCGGTGACCACTCATCATCTCAAGAAATTA 779
 Db 541 AAGATATCTAAGCTATTGAAACACTAGTATCGGTGACCACTCATCATCTCAAGAAATTA 600

QY 780 CTAGCTCAAGCAAAAGCATTTTAAACAAAACCCAGGCTATACGATTTATGAACGT 839
Db |||||
QY 601 CTAGCTCAAGCAAAAGCATTTTAAACAAAACCCAGGCTATACGATTTATGAACGT 660
Db |||||
QY 840 GACTCCTCAATCGTCACTGATGACATGACATTTTCGTCGATTTTACCAATGGATCAA 899
Db |||||
QY 661 GACTCCTCAATCGTCACTGATGACATGACATTTTCGTCGATTTTACCAATGGATCAA 720
Db |||||
QY 900 GAGTTTACTTACCGTGTAAATAATCGGGAACAAGCTTATAGATCAATAAAAAATCTGGT 959
Db |||||
QY 721 GAGTTTACTTACCATGTCAATAATCGGGAACAAGCTTATAGATCAATAAAAAATCTGGT 780
Db |||||
QY 960 CTGAATGAAGAAATAAACAACAACCTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAA 1019
Db |||||
QY 781 CTGAATGAAGAAATAAACAACAACCTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAA 840
Db |||||
QY 1020 GGGGAAAAGCCGTATGATCCCTTTGATCGGAGTCACCTTGAAACTGTTCACCATCAATAC 1079
Db |||||
QY 841 GGGGAAAAGCCGTATGATCCCTTTGATCGGAGTCACCTTGAAACTGTTCACCATCAATAC 900
Db |||||
QY 1080 GTTGATGTCGATACCAACGAATTGCTAAAGTGAGCAGCTCTTAAACAGCTAGCGACGT 1139
Db |||||
QY 901 GTTGATGTCGAACCAACGAATTGCTAAAGTGAGCAGCTCTTAAACAGCTAGCGACGT 960
Db |||||
QY 1140 AACTTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACAACAAT 1199
Db |||||
QY 961 AACTTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACAACAAT 1020
Db |||||
QY 1200 CTCGATGCTTTTGGTATTATGGACTATACCTTAACTGGAAAAGTAGAGGATAATCACGAT 1259
Db |||||
QY 1021 CTCGATGCTTTTGGTATTATGGACTATACCTTAACTGGAAAAGTAGAGGATAATCACGAT 1080
Db |||||
QY 1260 GACACCAACCGTATCATACCGTTTATATGGGCAAGCGA 1298
Db |||||
QY 1081 GACACCAACCGTATCATACCGTTTATATGGGCAAGCGA 1119
Db |||||

Search completed: February 1, 2006, 12:43:42
Job time : 294.256 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2006, 17:06:30 ; Search time 1372.44 Seconds
(without alignments)
10008.068 Million cell updates/sec

Title: US-09-940-235-10
Perfect score: 1661
Sequence: 1 gcaaccgccagcctagcc.....gaataagctgtaccatctaa 1661

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA_Main:*

- 1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:*
- 3: /cgn2_6/prodata/1/pubpna/US09A_PUBCOMB.seq:*
- 4: /cgn2_6/prodata/1/pubpna/US09B_PUBCOMB.seq:*
- 5: /cgn2_6/prodata/1/pubpna/US10A_PUBCOMB.seq:*
- 6: /cgn2_6/prodata/1/pubpna/US10B_PUBCOMB.seq:*
- 7: /cgn2_6/prodata/1/pubpna/US10C_PUBCOMB.seq:*
- 8: /cgn2_6/prodata/1/pubpna/US10D_PUBCOMB.seq:*
- 9: /cgn2_6/prodata/1/pubpna/US10E_PUBCOMB.seq:*
- 10: /cgn2_6/prodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1661	100.0	1661	3	US-09-940-235-10
2	1282.2	77.2	1377	3	US-09-940-235-5
3	1219.4	73.4	1327	3	US-09-940-235-6
4	1185	71.3	1541	3	US-09-940-235-9
5	1165.8	70.2	2096	3	US-09-940-235-12
6	1150.2	69.2	1782	3	US-09-940-235-11
7	1148.6	69.2	1245	3	US-09-940-235-1
8	995	59.9	1323	8	US-10-474-792-657
9	316.8	19.1	6988	7	US-10-236-392-1
10	316.8	19.1	7679	9	US-10-831-704-38
11	316.8	19.1	7680	5	US-10-171-311-63
12	316.8	19.1	7680	6	US-10-236-031B-69
13	316.8	19.1	7680	6	US-10-374-979-75
14	316.8	19.1	7680	7	US-10-182-936A-75
15	316.8	19.1	7680	7	US-10-641-643-1289
16	316.8	19.1	7680	7	US-10-717-597-222
17	316.8	19.1	7680	8	US-10-788-792-79
18	316.8	19.1	7680	8	US-10-477-238A-654
19	316.8	19.1	7680	8	US-10-680-287A-654
20	316.8	19.1	7680	8	US-10-278-698-88
21	316.8	19.1	7680	8	US-10-278-698-88
22	316.8	19.1	7680	8	US-10-278-698-603
23	316.8	19.1	7680	9	US-10-843-641A-5877

24	316.8	19.1	7680	9	US-10-477-173-654	Sequence 654, Appl
25	316.8	19.1	7680	9	US-10-852-335A-52	Sequence 52, Appl
26	316.8	19.1	7705	7	US-10-447-161-4	Sequence 4, Appl
27	316.8	19.1	8216	9	US-10-450-763-22270	Sequence 22270, A
28	316.4	19.0	777	3	US-09-940-235-3	Sequence 3, Appl
29	313.6	18.9	7795	5	US-10-084-817-2	Sequence 2, Appl
30	312	18.8	2127	6	US-10-210-120-49	Sequence 4288, Ap
31	312	18.8	2127	9	US-10-956-157-4288	Sequence 49, Appl
32	312	18.8	2127	9	US-10-909-035-49	Sequence 4288, Ap
33	312	18.8	3522	9	US-10-450-763-22266	Sequence 49, Appl
34	312	18.8	4295	6	US-10-144-194A-51	Sequence 22266, A
35	312	18.8	4295	8	US-10-491-566-51	Sequence 51, Appl
36	312	18.8	7361	7	US-10-236-392-3	Sequence 51, Appl
37	312	18.8	7677	9	US-10-956-157-4995	Sequence 3, Appl
38	312	18.8	7867	5	US-10-098-841-6	Sequence 4995, Ap
39	312	18.8	8027	7	US-10-447-161-8	Sequence 6, Appl
40	312	18.8	8027	9	US-10-734-564-27	Sequence 8, Appl
41	312	18.8	8027	9	US-10-852-335A-53	Sequence 27, Appl
42	312	18.8	8027	9	US-10-287-436A-81	Sequence 53, Appl
43	312	18.8	8062	5	US-10-098-841-5	Sequence 81, Appl
44	312	18.8	8137	5	US-10-098-841-8	Sequence 5, Appl
45	312	18.8	8230	5	US-10-098-841-7	Sequence 8, Appl

ALIGNMENTS

RESULT 1
US-09-940-235-10
; Sequence 10, Application US/09940235
; Publication No. US20030059921A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammar
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/09/940,235
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 1661
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hybrid cassette
US-09-940-235-10

Query Match	100.0%;	Score 1661;	DB 3;	Length 1661;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1661;	Conservative 0;	Mismatches 0;	Gaps 0;	
QY	1	GCAACCCCGCCAGCCTAGCCGGGTCCTCAACAGCAGGACGATCATGCGCACCCGCGG	60	
Db	1	GCAACCCCGCCAGCCTAGCCGGGTCCTCAACAGCAGGACGATCATGCGCACCCGCGG	60	
QY	61	CCAGGACCCCAACCGTCGCCGAGATCTCGATCCCGCGAAATTAATACGACTCATTATAGG	120	
Db	61	CCAGGACCCCAACCGTCGCCGAGATCTCGATCCCGCGAAATTAATACGACTCATTATAGG	120	
QY	121	AGACCACACCGGTTTCCTCTAGNAATATTTTCTTTAAGAGGAGATATACCA	180	


```

121 AGACCAACCGTTTCCCTCTAGAAATAATTTTGTATTAACTTTAAGAGGAGATATACCA 180
181 TGATTGCTGGACCTGAGTGGCTGTAGACGGTCCATCTCTCAACACGCAATTTGGTTG 240
181 TGATTGCTGGACCTGAGTGGCTGTAGACGGTCCATCTCTCAACACGCAATTTGGTTG 240
241 TTAGCGTTGCTGGTACTGTTGAGGGGACGAACTCAAGACATTAAGTCTTTAAATTTTTGAAA 300
241 TTAGCGTTGCTGGTACTGTTGAGGGGACGAACTCAAGACATTAAGTCTTTAAATTTTTGAAA 300
301 TCGATCTAACATCACGACTGCTCATGAGGAGGAGAGACAGCAAGGCTTTAAGTCCAAAT 360
301 TCGATCTAACATCACGACTGCTCATGAGGAGGAGAGACAGCAAGGCTTTAAGTCCAAAT 360
361 CAAAACCAATTTGCTACTGATAGTGGCGGATGTCTCATAAACCTTGAGAAAGCTGACTTAC 420
361 CAAAACCAATTTGCTACTGATAGTGGCGGATGTCTCATAAACCTTGAGAAAGCTGACTTAC 420
421 TAAAGGCTATTCAAGAAATTTGATCGCTAACTGCTCCACAGTAAACGACGACTACTTTGAGG 480
421 TAAAGGCTATTCAAGAAATTTGATCGCTAACTGCTCCACAGTAAACGACGACTACTTTGAGG 480
481 TCATTGATTTTGCAGCGATGCAACCATTTACTGATCGAAACGCGCAAGGCTCTACTTTGCTG 540
481 TCATTGATTTTGCAGCGATGCAACCATTTACTGATCGAAACGCGCAAGGCTCTACTTTGCTG 540
541 ACAAGATGGTTTCGGTAACTTTCGGTAACTTTCGGTAACTTTCGGTAACTTTCGGTAACTTTC 600
541 ACAAGATGGTTTCGGTAACTTTCGGTAACTTTCGGTAACTTTCGGTAACTTTCGGTAACTTTC 600
601 ATGTGCGGTTAGACCATATAGAAAGAAACCAATACAAACCAAGCGAAATCTGTTGATG 660
601 ATGTGCGGTTAGACCATATAGAAAGAAACCAATACAAACCAAGCGAAATCTGTTGATG 660
661 TGGATATACCTGTAAGTATTTTAACTTCTCCCTTAAACCTTGAACGATTCAGACCAAGGCTCTCA 720
661 TGGATATACCTGTAAGTATTTTAACTTCTCCCTTAAACCTTGAACGATTCAGACCAAGGCTCTCA 720
721 AAGATACCTAAGCTATTGAAACCACTAGCTATCGGTGACACCATCATCATCTCAAGAATTAC 780
721 AAGATACCTAAGCTATTGAAACCACTAGCTATCGGTGACACCATCATCATCTCAAGAATTAC 780
781 TAGCTCAGCACAAGCAATTTTAAACAAACCAACCAACCAAGGCTATAGGATTTAAGACGTG 840
781 TAGCTCAGCACAAGCAATTTTAAACAAACCAACCAACCAAGGCTATAGGATTTAAGACGTG 840
841 ACTCCTCAATCGTCACTCATGACAATGACATTTTCGGTACGATTTTACCAGTGGATCAAG 900
841 ACTCCTCAATCGTCACTCATGACAATGACATTTTCGGTACGATTTTACCAGTGGATCAAG 900
901 AGTTTACTTACCGTGTAAAAATCGGAAACAGCTTTATAGGATCAATAAAAAATCTGGTC 960
901 AGTTTACTTACCGTGTAAAAATCGGAAACAGCTTTATAGGATCAATAAAAAATCTGGTC 960
961 TGAATGAAGAAATAAACAACACTGATCTCTGAGAAATATTACGTCCTTTAAAAAG 1020
961 TGAATGAAGAAATAAACAACACTGATCTCTGAGAAATATTACGTCCTTTAAAAAG 1020
1021 GGGAAAGCGGTATGATCCCTTTGATCGCAGTCACTTTGAAACCTGTTACCACCAATACG 1080
1021 GGGAAAGCGGTATGATCCCTTTGATCGCAGTCACTTTGAAACCTGTTACCACCAATACG 1080
1081 TTGATGTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTTTAAGCTAGCGAAGCTA 1140
1081 TTGATGTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTTTAAGCTAGCGAAGCTA 1140
1141 ACTTAGACTTCAGAGATTTATAGATCTCTCGTGATAAGGCTAAATCTAACAACAATC 1200
1141 ACTTAGACTTCAGAGATTTATAGATCTCTCGTGATAAGGCTAAATCTAACAACAATC 1200
1201 TCGATGCTTTTGGTATTGACTATACCTTAACCTGGAAGCTAGAGGATATACGATG 1260

```

```

Db 1201 TCGATGCTTTTGGTATTATGGACTATACCTTAACTGGAAGTAGAGATAATCACGATG 1260
Qy 1261 ACACCAACCGTATCATAAACCGTTTATATGCGCAAGCCGACCGAGGAGAGATCTAGCT 1320
Db 1261 ACACCAACCGTATCATAAACCGTTTATATGCGCAAGCCGACCGAGGAGAGATCTAGCT 1320
Qy 1321 ATCATTTAGCCGCTGGTGGTTCAGGCGCAGCAAAATGGTTTCAGCCCGCTCCCGGTGGCTG 1380
Db 1321 ATCATTTAGCCGCTGGTGGTTCAGGCGCAGCAAAATGGTTTCAGCCCGCTCCCGGTGGCTG 1380
Qy 1381 TCAGTCAAGCAAGCCCGGTTGTTATGACAAATGGTTCAGGCGCAGCAAAATGGTTCAGGCG 1440
Db 1381 TCAGTCAAGCAAGCCCGGTTGTTATGACAAATGGTTCAGGCGCAGCAAAATGGTTCAGGCG 1440
Qy 1441 GGGAGCGGACCTTACCTAGGTAATGTGTTGGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1500
Db 1441 GGGAGCGGACCTTACCTAGGTAATGTGTTGGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1500
Qy 1501 TTAACTGCGAAAGCTTAAACCTGGAAGCTGGAAGACTTCTTTTGACAAGTACACTGGGAACA 1560
Db 1501 TTAACTGCGAAAGCTTAAACCTGGAAGCTGGAAGACTTCTTTTGACAAGTACACTGGGAACA 1560
Qy 1561 CTTACCGAGTGGGTGACACTTATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACCT 1620
Db 1561 CTTACCGAGTGGGTGACACTTATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACCT 1620
Qy 1621 GCATCGGGCTGGGCGAGGAGAAATAGCTGTACCATCTAA 1661
Db 1621 GCATCGGGCTGGGCGAGGAGAAATAGCTGTACCATCTAA 1661

```

```

RESULT 2
US-09-940-235-5
; Sequence 5, Application US/09940235
; Publication No. US20030059921A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammarra
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/09/940,235
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1377
; TYPE: DNA
; ORGANISM: Streptococcus equisimilis
US-09-940-235-5

```

```

Query Match 77.2%; Score 1282.2; DB 3; Length 1377;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1284; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 51 GCACCCGTGGCAGGACCCAAACGCTGCCGAGATCTCGATCCGCGGAAATTAATACCACT 110
Db 1 GCACCCGTGGCAGGACCCAAACGCTGCCGAGATCTCGATCCGCGGAAATTAATACCACT 60
Qy 111 CACTATAGGAGACCAACGCTTCCCTCTAGAAATAATTTGTTTAACTTTAAGAAGG 170
Db 61 CACTATAGGAGACCAACGCTTCCCTCTAGAAATAATTTGTTTAACTTTAAGAAGG 120

```


Db 421 CGGAAGGTCTACTTTGCTGCAAAAGATGGTTTCGGTAACTTGGCGACCAACTGTCCA 480
QY 581 AGAATTTTGTCTAGCGGACATGTGCGGTGAGCATATATAAGAAACCAATACAAA 640
Db 481 AGAATTTTGTCTAGCGGACATGTGCGGTGAGCATATATAAGAAACCAATACAAA 540
QY 641 CCAAGCGAAATCTGTGTGATGTGAATATATCTGACAGTTTACTCCCTTAAACCTGATGA 700
Db 541 CCAAGCGAAATCTGTGTGATGTGAATATATCTGACAGTTTACTCCCTTAAACCTGATGA 600
QY 701 CGATTTACAGCCAGGTCTCAAAGATATTAAGCTATTGAAAAACATAGCTATCGGTGACAC 760
Db 601 CGATTTACAGCCAGGTCTCAAAGATATTAAGCTATTGAAAAACATAGCTATCGGTGACAC 660
QY 761 CATCATCTCAAGAAATTAAGCTCAAGCAACAAAGCAATTTAAACAAAACCAACCCAGG 820
Db 661 CATCATCTCAAGAAATTAAGCTCAAGCAACAAAGCAATTTAAACAAAACCAACCCAGG 720
QY 821 CTATACGATTTATGAACGTGACTCTCAATCGTCACTCATGACAATGACATTTTCGGTAC 880
Db 721 CTATACGATTTATGAACGTGACTCTCAATCGTCACTCATGACAATGACATTTTCGGTAC 780
QY 881 GATTTTACCAATGGAATCAAGAGTTTACTTACCGTGTAAATAATCGGAAACAAAGCTTATAG 940
Db 781 GATTTTACCAATGGAATCAAGAGTTTACTTACCGTGTAAATAATCGGAAACAAAGCTTATAG 840
QY 941 GATCAATAAATAATCTGTGATGAATGAAGAAATAACAACTGACCTGATCTGTAGAA 1000
Db 841 GATCAATAAATAATCTGTGATGAATGAAGAAATAACAACTGACCTGATCTGTAGAA 900
QY 1001 ATATTAGCTCTTAAAGAGGGGAAAGCGGTATGATCCCTTTGATCGGAGTCACTTGA 1060
Db 901 ATATTAGCTCTTAAAGAGGGGAAAGCGGTATGATCCCTTTGATCGGAGTCACTTGA 960
QY 1061 ACTGTTCAACCATCAATACGTTGATGTCGATACCAACCAATGCTTAAAGAGTGAGCAGCT 1120
Db 961 ACTGTTCAACCATCAATACGTTGATGTCGATACCAACCAATGCTTAAAGAGTGAGCAGCT 1020
QY 1121 CTTAAAGCTAGCGAAACGTAACCTTAGAGCTTCAGAGATTATACGATTCCTCGTAAAGGC 1180
Db 1021 CTTAAAGCTAGCGAAACGTAACCTTAGAGCTTCAGAGATTATACGATTCCTCGTAAAGGC 1080
QY 1181 TAAACTACTTACACCAATCTCGATGCTTTTGGTATTATGACTATACCTTAACTGAAA 1240
Db 1081 TAAACTACTTACACCAATCTCGATGCTTTTGGTATTATGACTATACCTTAACTGAAA 1140
QY 1241 AGTAGAGATTAATCAAGATGACCAACCGGTATCATACCGTTTATATGGGCAAGCGAC 1300
Db 1141 AGTAGAGATTAATCAAGATGACCAACCGGTATCATACCGTTTATATGGGCAAGCGAC 1200
QY 1301 CGAAGGAGAGATGCTAGCTATCATTTTAGCCGGTGGT 1337
Db 1201 CGAAGGAGAGATGCTAGCTATCATTTTAGCCCTATGAT 1237

RESULT 4
US-09-940-235-9
; Sequence 9, Application US/09940235
; Publication No. US2003005921A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammara
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002

; CURRENT APPLICATION NUMBER: US/09/940,235
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1541
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hybrid cassette
US-09-940-235-9

Query Match 71.3%; Score 1185; DB 3; Length 1541;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 1194; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 151 TTGTTTAACTTTAAGAGGAGATATACCATGATTTGCTGGACCTGAGTGTGCTAGACC 210
Db 1 TTTGTTTAACTTTAAGAGGAGATATACCATGATTTGCTGGACCTGAGTGTGCTAGATC 60
QY 211 GTCCATCTGTCAACAACAGCCCAATTTGTTGTTAGCGTTGCTGTCTGTTGAGGAGCA 270
Db 61 GTCCCTCTCTAAATAACAGCCCAATTTGTTGTTAGCGTTGCTGTCTGTTGAGGAGCA 120
QY 271 ATCAAGACATTAAGTCTTAAATTTTGAATTCATCTAAACATCAACGACCTGCTCATGG 330
Db 121 ATCAAGACATTAAGTCTTAAATTTTGAATTCATCTAAACATCAACGACCTGCTCATGG 180
QY 331 GAAAGACAGAGCAAGGCTTAAAGTCCAAATCAAAACCAATTTGCTGTCTGATGAGGCGCA 390
Db 181 GAAAGACAGAGCAAGGCTTAAAGTCCAAATCAAAACCAATTTGCTGTCTGATGAGGCGCA 240
QY 391 TGTCAATAACTTGAAGAGCTGACTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCT 450
Db 241 TGTCAATAACTTGAAGAGCTGACTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCT 300
QY 451 AGCTCCACAGTAAACGACGACTTCTTTGAGGTCAATTCATTTTGAAGCGATGCAACCAATTA 510
Db 301 AGCTCCACAGTAAACGACGACTTCTTTGAGGTCAATTCATTTTGAAGCGATGCAACCAATTA 360
QY 511 CTGATCGAAACGCAAGGCTTACTTTGCTGACAAAGATGTTGCGTAACTTTCGCGACCC 570
Db 361 CTGATCGAAACGCAAGGCTTACTTTGCTGACAAAGATGTTGCGTAACTTTCGCGACCC 420
QY 571 AACCTGTCCAAGAAATTTTGTCTAAGCGGACATGTGCGGTTAGACCATATATAAGAAAC 630
Db 421 AACCTGTCCAAGAAATTTTGTCTAAGCGGACATGTGCGGTTAGACCATATATAAGAAAC 480
QY 631 CAATACAAAACCAAGCGAAATCTGTTGATGTGGAATATATCTGTACAGTTTACTCCCTTAA 690
Db 481 CAATACAAAACCAAGCGAAATCTGTTGATGTGGAATATATCTGTACAGTTTACTCCCTTAA 540
QY 691 ACCCTGATGATCGAATTTTCAAGCCAGGCTCTCAAGATTAAGCTTAAAGCAACTAGCTA 750
Db 541 ACCCTGATGATCGAATTTTCAAGCCAGGCTCTCAAGATTAAGCTTAAAGCAACTAGCTA 600
QY 751 TCGGTGACACCATCATCTCAAGATTAAGCTTCAAGCAACAAAGCATTTTAAACAAA 810
Db 601 TCGGTGACACCATCATCTCAAGATTAAGCTTCAAGCAACAAAGCATTTTAAACAAA 660
QY 811 ACCACCCAGGCTATACGATTTTGAACGAGTCCCTCAATCGTCACTCATGACAAATGACA 870
Db 661 ACCACCCAGGCTATACGATTTTGAACGAGTCCCTCAATCGTCACTCATGACAAATGACA 720
QY 871 TTTTCCGTACGATTTTACCAATGGATCAAGAGTTTACTTACCGTGTGTTAAATAATCGGAC 930
Db 721 TTTTCCGTACGATTTTACCAATGGATCAAGAGTTTACTTACCGTGTGTTAAATAATCGGAC 780
QY 931 AAGCTTATAGGATCAATATAAAAAAATCTGGTCTGAATGAAGAAATAAACAACACTGACCTGA 990

[illegible]

RESULT 5

```

US-09-940-235-12
; Sequence 12, Application US/09940235
; Publication No. US20030059921A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chaic
; APPLICANT: Rajagopal, Kammara
; APPLICANT: Nihalani, Deepak
; APPLICANT: Nuhdarani, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/09/940,235
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2096
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hybrid cassette
US-09-940-235-12

```

	Query Match	70.2%	Score 1165.8	DB 3	Length 2096	
	Best Local Similarity	99.4%	Pred. No. 0			
	Matches 1170	Conservative	0	Mismatches 7	Indels 0	Gaps 0
Qy	183	ATTGCTGGACCTGAGTGCTGCTAGACCGTCCATCTGTCAACACGACCAATTGGTTGTT	242			
Db	588	ATTGCTGGACCTGAGTGCTGCTAGACCGTCCATCTGTCAACACGACCAATTGGTTGTT	647			

Qy	243	AGCGTTCGTGTA	CTGTTGAGGGGACGAATCAAGACATTA	GTCTTTAAATTTTTTTGAAATC	302
Db	648	AGCGTTCGTGTA	CTGTTGAGGGGACGAATCAAGACATTA	GTCTTTAAATTTTTTTGAAATC	707
Qy	303	GATCTAACATC	ACGACCTGCTCATGAGGAAAGACAGAGCAAGC	TTAAGTCCAAATCA	362
Db	708	GATCTAACATC	ACGACCTGCTCATGAGGAAAGACAGAGCAAGC	TTAAGTCCAAATCA	767
Qy	363	AAACCAATTGCT	ACTGATAGTGGGCGGATGTCA	CATAAACTTTGAGAAAGCTGACTTACTA	422
Db	768	AAACCAATTGCT	ACTGATAGTGGGCGGATGTCA	CATAAACTTTGAGAAAGCTGACTTACTA	827
Qy	423	AAGGCTATTCA	AGAAACAAATTAATGCTGTAA	CGTCCACAGTAAACGACACTCTTTGAGGTC	482
Db	828	AAGGCTATTCA	AGAAACAAATTAATGCTGTAA	CGTCCACAGTAAACGACACTCTTTGAGGTC	887
Qy	483	ATTGATTTTGC	AAAGCGATGCAACCAATTA	CTGATCGAAACGCAAGGTCCTCTTTGCTGAC	542
Db	888	ATTGATTTTGC	AAAGCGATGCAACCAATTA	CTGATCGAAACGCAAGGTCCTCTTTGCTGAC	947
Qy	543	AAAGATGGTTCC	GTAAACCTTCCCGACCCAA	CCCTGTCCAAGAAATTTTGTCTAAGCGGACAT	602
Db	948	AAAGATGGTTCC	GTAAACCTTCCCGACCCAA	CCCTGTCCAAGAAATTTTGTCTAAGCGGACAT	1007
Qy	603	GTGGCGGTTAGA	CCATATAAAGAAAAACCAATA	CAAAACCAAGCGAAATCTGTTGATGTG	662
Db	1008	GTGGCGGTTAGA	CCATATAAAGAAAAACCAATA	CAAAACCAAGCGAAATCTGTTGATGTG	1067
Qy	663	GAATATACTGT	PACAGTTTACTCCCTTAA	ACCTGATGACGATTTTCAGACAGGTCCTCAA	722
Db	1068	GAATATACTGT	PACAGTTTACTCCCTTAA	ACCTGATGACGATTTTCAGACAGGTCCTCAA	1127
Qy	723	GATACTAAGCT	TAATTGAAAAACACTAGCTAT	CGGTGACACCATCACATCTCAAGAATTA	782
Db	1128	GATACTAAGCT	TAATTGAAAAACACTAGCTAT	CGGTGACACCATCACATCTCAAGAATTA	1187
Qy	783	GCTCAAGCA	CAAAAGCAATTTTAAACAAAA	CCACCAGGCTATACGATTTATGAACGTCAC	842
Db	1188	GCTCAAGCA	CAAAAGCAATTTTAAACAAAA	CCACCAGGCTATACGATTTATGAACGTCAC	1247
Qy	843	TCCTCAATCGT	CACATGACAAATGACATTTTCCGTAGGATTTT	ACCAATGGATCAAGAG	902
Db	1248	TCCTCAATCGT	CACATGACAAATGACATTTTCCGTAGGATTTT	ACCAATGGATCAAGAG	1307
Qy	903	TTTACTTACCGT	GTFTTAAAAATCGGGAA	CAAGCTTATAGGATCAATAAAAAATCTGGCTG	962
Db	1308	TTTACTTACCGT	GTFTTAAAAATCGGGAA	CAAGCTTATAGGATCAATAAAAAATCTGGCTG	1367
Qy	963	AATGAAGAAAT	TAACACACCTGATCTCTGAGAAATATTA	ACGTCCTTTAAAAAAGG	1022
Db	1368	AATGAAGAAAT	TAACACACCTGATCTCTGAGAAATATTA	ACGTCCTTTAAAAAAGG	1427
Qy	1023	GAAGAAGCGT	ATGATCCCTTTGATCGCAGTCA	TTTGAAACGTGTTACCATCAATACGTT	1082
Db	1428	GAAGAAGCGT	ATGATCCCTTTGATCGCAGTCA	TTTGAAACGTGTTACCATCAATACGTT	1487
Qy	1083	GATGTCGAT	CAACCAAGTAATGCTTAAAAAGTGAGCAGCT	CTTTAACAGCTAGCGAACGTAAC	1142
Db	1488	GATGTCGAT	CAACCAAGTAATGCTTAAAAAGTGAGCAGCT	CTTTAACAGCTAGCGAACGTAAC	1547
Qy	1143	TTAGACTTCAG	AGATTTTATACGATCCTCGTGATAAGGCT	AAATCTCTACAAATCTC	1202
Db	1548	TTAGACTTCAG	AGATTTTATACGATCCTCGTGATAAGGCT	AAATCTCTACAAATCTC	1607
Qy	1203	GATGCTTTTCG	TATATGGA	CTATACTTTAACTGGAAGAGTACAGGATTAATCA	1262
Db	1608	GATGCTTTTCG	TATATGGA	CTATACTTTAACTGGAAGAGTACAGGATTAATCA	1667
Qy	1263	ACCAACCGT	ATCATAAACCGTTTATATG	GGCAAGCCGCAAGGAGAGATGCTAGCTAT	1322
Db	1668	ACCAACCGT	ATCATAAACCGTTTATATG	GGCAAGCCGCAAGGAGAGATGCTAGCTAT	1727
Qy	1323	CATTTAGCCG	TGGTGTGAGGGCGACAAATGGTTC	1359	


```

Db 439 AAGATGTTGGTAAACCTTGCGGACCCCAACCTGTCACAGAAATTTTGTGTTAAAGGACAT 498
Qy 603 GTGCGCGTTAGACCATATAAGAAAAACCAATACAAAAACCAAGCGAAATCTGTTGATGTG 662
Db 499 GTGCGCGTTAGACCATATAAGAAAAACCAAGTACAAAAATCAGCAAAATCTGTTGATGTA 558
Qy 663 GAATATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGTCTCTCAA 722
Db 559 GAATATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGTCTCTCAA 618
Qy 723 GATACATAGCTATTGAAACACATAGCTATCGGTGACACCATCATCTCAAGAAATTAATA 782
Db 619 GATACATAGCTATTGAAACACATAGCTATCGGTGACACCATCATCTCAAGAAATTAATA 678
Qy 783 GCTCAAGCACAAAGCAATTTTAAACAAAAACCAAGGCTATACGATTTTATGAACTGAC 842
Db 679 GCTCAAGCACAAAGCAATTTTAAACAAAAACCAAGGCTATACGATTTTATGAACTGAC 738
Qy 843 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 902
Db 739 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 798
Qy 903 TTTACTTACCGTTTAAAAATCGGGAACAAAGCTTATAGGATCAATAAAAAATCTGTTCTG 962
Db 799 TTTACTTACCATGTCAAAAATCGGGAACAAAGCTTATGAGATCAATCTTAAAAACAGGTATT 858
Qy 963 AATGAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 1022
Db 859 AAGAAAAAACAAGCAACTGATCTGCTGCTGAGAAATATTACGTCCTTAAAAAAGGG 918
Qy 1023 GAAAAAGCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTTACCAATCAATACGTT 1082
Db 919 GAAAAAGCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTTACCAATCAATACGTT 978
Qy 1083 GATGTCGATACCAACGAATGCTTAAAGTAGACAGCTCTTAAACAGCTAGCGAACGTAAC 1142
Db 979 GATGTCGATACCAACGAATGCTTAAAGTAGACAGCTCTTAAACAGCTAGCGAACGTAAC 1038
Qy 1143 TTAGACTTCAGAGATTATACGATCCTCGTGATAGAGCTTAAACTACTCTACAACTCTC 1202
Db 1039 TTAGACTTCAGAGATTATACGATCCTCGTGATAGAGCTTAAACTACTCTACAACTCTC 1098
Qy 1203 GATGCTTTGTTGTTATGATGATATACCTTAACTGGAAGAGTAGAGGATATACGATGAC 1262
Db 1099 GATGCTTTGTTGTTATGATGATATACCTTAACTGGAAGAGTAGAGGATATACGATGAC 1158
Qy 1263 ACCAACCGTATCAACCGTTTATATGGGCAAGCGACCCGAGAGAGAGATGCTAGCTAT 1322
Db 1159 AATAATCGTGTGTTACAGTTTATATGGGCAAGCGCCCTTAAAGGGCAAGGCTAGCTAT 1218
Qy 1323 CATTTAGCGGTGGT 1337
Db 1219 CATTTAGCTTATGAT 1233

```

RESULT 9
US-10-236-392-1

; Sequence 1, Application US/10236392
; Publication No. US20040067490A1
; GENERAL INFORMATION:

; APPLICANT: Anderson, David W
; APPLICANT: Boldog, Ferenc L
; APPLICANT: Burgess, Catherine, E
; APPLICANT: Casman, Stacie J
; APPLICANT: Catterton, Elina
; APPLICANT: Chapoval, Andrei
; APPLICANT: Crabtree, Julie
; APPLICANT: Edinger, Shlomit, R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gerlach, Valerie
; APPLICANT: Gorman, Linda
; APPLICANT: Grosse, William M
; APPLICANT: Gusev, Vladamir

```

; APPLICANT: Kekuda, Ramesh
; APPLICANT: LaRoche, William J
; APPLICANT: Li, Li
; APPLICANT: MacDougall, John R
; APPLICANT: Maliyankar, Uriel M
; APPLICANT: Miller, Charles E
; APPLICANT: Millet, Isabelle
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Patturajan, Meera
; APPLICANT: Pena, Carol A
; APPLICANT: Peyman, John A
; APPLICANT: Rastelli, Luca
; APPLICANT: Reiger, Daniel K
; APPLICANT: Rothenberg, Mark E
; APPLICANT: Shenoy, Suresh
; APPLICANT: Shimkets, Richard A
; APPLICANT: Smithson, Glennda
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-442A
; CURRENT APPLICATION NUMBER: US/10/236,392
; CURRENT FILING DATE: 2002-09-06
; PRIOR APPLICATION NUMBER: US09/540,763
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: US60/390,155
; PRIOR FILING DATE: 2002-06-19
; PRIOR APPLICATION NUMBER: US09/635,949
; PRIOR FILING DATE: 2000-08-10
; PRIOR APPLICATION NUMBER: US60/318,765
; PRIOR FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: US60/357,303
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US60/367,753
; PRIOR FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER: US60/369,479
; PRIOR FILING DATE: 2002-04-02
; PRIOR APPLICATION NUMBER: US09/659,634
; PRIOR FILING DATE: 2000-09-12
; PRIOR APPLICATION NUMBER: US60/318,120
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US60/318,130
; PRIOR FILING DATE: 2001-09-07
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 794
; SOFTWARE: Custom
; SEQ ID NO 1
; LENGTH: 6988
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (26)..(6986)
; US-10-236-392-1

```

Query Match 19.1%; Score 316.8; DB 7; Length 6988;
Best Local Similarity 99.4%; Pred. No. 4e-79;
Matches 318; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

Qy 1339 GTCAGGCGCAGCAAAATGGTTTCAGCCCCCAGTCCCGGTGGTGTGTCAGTCAAGCAAGCCG 1398
Db 12 GGCAGGCTCAGCAAAATGGTTTCAGCCCCCAGTCCCGGTGGTGTGTCAGTCAAGCAAGCCG 71
Qy 1399 GTTGTATGCAATGGAACACATATCAGATAATCAACAGTGGGAGGACCTACCTAG 1458
Db 72 GTTGTATGCAATGGAACACATATCAGATAATCAACAGTGGGAGGACCTACCTAG 131
Qy 1459 GTAAATGTTGTTTGTACTTGTATCGAGAACCCAGGTTTTTAACCTGCCAAAGTAAAC 1518
Db 132 GTAAATGTTGTTTGTACTTGTATCGAGAACCCAGGTTTTTAACCTGCCAAAGTAAAC 191
Qy 1519 CTGAAGCTGAAGAGACTTGTCTTTGACAAGTACACTCGGGAACACTTACCGAGTGGGTGACA 1578
Db 192 CTGAAGCTGAAGAGACTTGTCTTTGACAAGTACACTCGGGAACACTTACCGAGTGGGTGACA 251

```



```
QY 1579 CTTATGAGGCTCTAAAGACTCATGATCTGGAGCTGTACCTGCATCGGGCTGGGCGAG 1638
Db 252 CTTATGAGGCTCTAAAGACTCATGATCTGGAGCTGTACCTGCATCGGGCTGGGCGAG 311
QY 1639 GGAGATAAGCTGTACCATC 1658
Db 312 GGAGATAAGCTGTACCATC 331

RESULT 10
US-10-831-704-38
; Sequence 38, Application US/10831704
; Publication No. US20050100931A1
; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND
; TITLE OF INVENTION: ASSESSMENT
; TITLE OF INVENTION: OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 07334-074001
; CURRENT APPLICATION NUMBER: US/10/831,704
; CURRENT FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US/10/155,653
; PRIOR FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US/09/220,132
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: US 60/079,303
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: US 60/068,821
; PRIOR FILING DATE: 1997-12-24
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 7679
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-831-704-38

Query Match 19.1%; Score 316.8; DB 9; Length 7679;
Best Local Similarity 99.4%; Pred. No. 4.2e-79;
Matches 318; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1339 GTCAGGCGCAGCAAAATGTTTCAGCCCGAGTCCCGGTGGCTGTCAAGTCAAGCAAGCCCG 1398
Db 12 GGCAGGCTCAGCAATGTTTCAGCCCGAGTCCCGGTGGCTGTCAAGTCAAGCAAGCCCG 71
QY 1399 GTTGTATGACATCGAAACACTATCAGATAAATCAACAGTGGAGCGGACCTACCTAG 1458
Db 72 GTTGTATGACATCGAAACACTATCAGATAAATCAACAGTGGAGCGGACCTACCTAG 131
QY 1459 GTAATGTTGTTGTTTGTACTTGTATGGAGGAAGCCGAGGTTTAACTCGGAAAGTAAAC 1518
Db 132 GTAATGTTGTTGTTTGTACTTGTATGGAGGAAGCCGAGGTTTAACTCGGAAAGTAAAC 191
QY 1519 CTGAAGCTGAAGAGACTTGTCTTTGACAAAGTACACTGGGAACACTTACCGAGTGGTGACA 1578
Db 192 CTGAAGCTGAAGAGACTTGTCTTTGACAAAGTACACTGGGAACACTTACCGAGTGGTGACA 251
QY 1579 CTTATGAGGCTCTAAAGACTCCATGATCTGGAGACTGTACTGATCGGGCTGGGCGAG 1638
Db 252 CTTATGAGGCTCTAAAGACTCCATGATCTGGAGACTGTACTGATCGGGCTGGGCGAG 311
QY 1639 GGAGATAAGCTGTACCATC 1658
Db 312 GGAGATAAGCTGTACCATC 331

RESULT 11
US-10-831-704-38
; Sequence 38, Application US/10831704
; Publication No. US20050100931A1
; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND
; TITLE OF INVENTION: ASSESSMENT
; TITLE OF INVENTION: OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 07334-074001
; CURRENT APPLICATION NUMBER: US/10/831,704
; CURRENT FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US/10/155,653
; PRIOR FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US/09/220,132
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: US 60/079,303
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: US 60/068,821
; PRIOR FILING DATE: 1997-12-24
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 7679
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-831-704-38

Query Match 19.1%; Score 316.8; DB 9; Length 7679;
Best Local Similarity 99.4%; Pred. No. 4.2e-79;
Matches 318; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1339 GTCAGGCGCAGCAAAATGTTTCAGCCCGAGTCCCGGTGGCTGTCAAGTCAAGCAAGCCCG 1398
Db 12 GGCAGGCTCAGCAATGTTTCAGCCCGAGTCCCGGTGGCTGTCAAGTCAAGCAAGCCCG 71
QY 1399 GTTGTATGACATCGAAACACTATCAGATAAATCAACAGTGGAGCGGACCTACCTAG 1458
Db 72 GTTGTATGACATCGAAACACTATCAGATAAATCAACAGTGGAGCGGACCTACCTAG 131
QY 1459 GTAATGTTGTTGTTTGTACTTGTATGGAGGAAGCCGAGGTTTAACTCGGAAAGTAAAC 1518
Db 132 GTAATGTTGTTGTTTGTACTTGTATGGAGGAAGCCGAGGTTTAACTCGGAAAGTAAAC 191
QY 1519 CTGAAGCTGAAGAGACTTGTCTTTGACAAAGTACACTGGGAACACTTACCGAGTGGTGACA 1578
Db 192 CTGAAGCTGAAGAGACTTGTCTTTGACAAAGTACACTGGGAACACTTACCGAGTGGTGACA 251
QY 1579 CTTATGAGGCTCTAAAGACTCCATGATCTGGAGACTGTACTGATCGGGCTGGGCGAG 1638
Db 252 CTTATGAGGCTCTAAAGACTCCATGATCTGGAGACTGTACTGATCGGGCTGGGCGAG 311
QY 1639 GGAGATAAGCTGTACCATC 1658
Db 312 GGAGATAAGCTGTACCATC 331

RESULT 11
US-09-964-824A-574
; Sequence 574, Application US/09964824A
; Patent No. US20020102531A1
; GENERAL INFORMATION:
; APPLICANT: Horrihan, Stephen
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
```

```
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-73
; CURRENT APPLICATION NUMBER: US/09/964,824A
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/60/236,033
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,032
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,028
; PRIOR FILING DATE: 2000-09-28
; NUMBER OF SEQ ID NOS: 583
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 574
; LENGTH: 7680
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-964-824A-574

Query Match 19.1%; Score 316.8; DB 3; Length 7680;
Best Local Similarity 99.4%; Pred. No. 4.2e-79;
Matches 318; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1339 GTCAGGCGCAGCAAAATGTTTCAGCCCGAGTCCCGGTGGCTGTCAAGTCAAGCAAGCCCG 1398
Db 12 GGCAGGCTCAGCAATGTTTCAGCCCGAGTCCCGGTGGCTGTCAAGTCAAGCAAGCCCG 71
QY 1399 GTTGTATGACATCGAAACACTATCAGATAAATCAACAGTGGAGCGGACCTACCTAG 1458
Db 72 GTTGTATGACATCGAAACACTATCAGATAAATCAACAGTGGAGCGGACCTACCTAG 131
QY 1459 GTAATGTTGTTGTTTGTACTTGTATGGAGGAAGCCGAGGTTTAACTCGGAAAGTAAAC 1518
Db 132 GTAATGTTGTTGTTTGTACTTGTATGGAGGAAGCCGAGGTTTAACTCGGAAAGTAAAC 191
QY 1519 CTGAAGCTGAAGAGACTTGTCTTTGACAAAGTACACTGGGAACACTTACCGAGTGGTGACA 1578
Db 192 CTGAAGCTGAAGAGACTTGTCTTTGACAAAGTACACTGGGAACACTTACCGAGTGGTGACA 251
QY 1579 CTTATGAGGCTCTAAAGACTCCATGATCTGGAGACTGTACTGATCGGGCTGGGCGAG 1638
Db 252 CTTATGAGGCTCTAAAGACTCCATGATCTGGAGACTGTACTGATCGGGCTGGGCGAG 311
QY 1639 GGAGATAAGCTGTACCATC 1658
Db 312 GGAGATAAGCTGTACCATC 331

RESULT 12
US-10-171-311-63
; Sequence 63, Application US/10171311
; Publication No. US20030087270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Glatt, Karen
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Hoerhah, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
```



```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63
; LENGTH: 7680
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-171-311-63

Query Match          19.1%; Score 316.8; DB 5; Length 7680;
Best Local Similarity 99.4%; Pred. No. 4.2e-79;
Matches 318; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1339 GTCAGGCGCAGCAATGGTTTCAGCCCGCAGTCCCGGTGCTGTCAAGCAAGCCCG 1398
Db 12 GGCAGGCTCAGCAATGGTTTCAGCCCGCAGTCCCGGTGCTGTCAAGCAAGCCCG 71

Qy 1399 GTTGTATGACAAATGGAAACACTATCAGATAAATCAACAGTGGGAGCGGACCTACCTAG 1458
Db 72 GTTGTATGACAAATGGAAACACTATCAGATAAATCAACAGTGGGAGCGGACCTACCTAG 131

Qy 1459 GTAATGTCTGGTTGTACTTGTATGGAGGAAGCCGAGGTTTTAACTCGAAAGTAAAC 1518
Db 132 GTAATGTCTGGTTGTACTTGTATGGAGGAAGCCGAGGTTTTAACTCGAAAGTAAAC 191

Qy 1519 CTGAAGCTGAAGAGACTTGTCTTTGACAAAGTACACTGGGAAACACTTACCGAGTGGGTGACA 1578
Db 192 CTGAAGCTGAAGAGACTTGTCTTTGACAAAGTACACTGGGAAACACTTACCGAGTGGGTGACA 251

Qy 1579 CTTATGAGCGTCCCTAAAGACTCCATGATCTGGGACTGTACTCGCATCGGGGCTGGGCGAG 1638
Db 252 CTTATGAGCGTCCCTAAAGACTCCATGATCTGGGACTGTACTCGCATCGGGGCTGGGCGAG 311

Qy 1639 GGAGATAAAGCTGTACCATC 1658
Db 312 GGAGATAAAGCTGTACCATC 331

RESULT 14
US-10-374-979-75
; Sequence 75, Application US/10374979
; Publication No. US20030219793A1
; GENERAL INFORMATION:
; APPLICANT: John P. Carulli et al.
; TITLE OF INVENTION: THE HIGH BONE MASS GENE OF 11q13.3
; FILE REFERENCE: 032796-021
; CURRENT APPLICATION NUMBER: US/10/374,979
; CURRENT FILING DATE: 2003-03-04
; PRIOR APPLICATION NUMBER: US 09/544,398
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 09/543,771
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 09/229,319
; PRIOR FILING DATE: 1999-01-13
; PRIOR APPLICATION NUMBER: US 60/071,449
; PRIOR FILING DATE: 1998-01-13
; PRIOR APPLICATION NUMBER: US 60/105,511
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 109
; SEQ ID NO 75
; LENGTH: 7680
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-374-979-75

Query Match          19.1%; Score 316.8; DB 6; Length 7680;
Best Local Similarity 99.4%; Pred. No. 4.2e-79;
Matches 318; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1339 GTCAGGCGCAGCAAAATGGTTTCAGCCCGCAGTCCCGGTGCTGTCAAGCAAGCCCG 1398
Db 12 GGCAGGCTCAGCAAAATGGTTTCAGCCCGCAGTCCCGGTGCTGTCAAGCAAGCCCG 71

Qy 1399 GTTGTATGACAAATGGAAACACTATCAGATAAATCAACAGTGGGAGCGGACCTACCTAG 1458
Db 72 GTTGTATGACAAATGGAAACACTATCAGATAAATCAACAGTGGGAGCGGACCTACCTAG 131

Qy 1459 GTAATGTCTGGTTGTACTTGTATGGAGGAAGCCGAGGTTTTAACTCGAAAGTAAAC 1518
Db 132 GTAATGTCTGGTTGTACTTGTATGGAGGAAGCCGAGGTTTTAACTCGAAAGTAAAC 191

Qy 1519 CTGAAGCTGAAGAGACTTGTCTTTGACAAAGTACACTGGGAAACACTTACCGAGTGGGTGACA 1578
Db 192 CTGAAGCTGAAGAGACTTGTCTTTGACAAAGTACACTGGGAAACACTTACCGAGTGGGTGACA 251

Qy 1579 CTTATGAGCGTCCCTAAAGACTCCATGATCTGGGACTGTACTCGCATCGGGGCTGGGCGAG 1638
Db 252 CTTATGAGCGTCCCTAAAGACTCCATGATCTGGGACTGTACTCGCATCGGGGCTGGGCGAG 311

Qy 1639 GGAGATAAAGCTGTACCATC 1658
Db 312 GGAGATAAAGCTGTACCATC 331

RESULT 13
US-10-236-031B-69
; Sequence 69, Application US/10236031B
; Publication No. US20030219760A1
; GENERAL INFORMATION:
; APPLICANT: Gordon, Gavin J.
; APPLICANT: Jensen, Roderick V.
; APPLICANT: Gullans, Steven R.
; APPLICANT: Bueno, Raphael
; TITLE OF INVENTION: Diagnostic and Prognostic Tests
; FILE REFERENCE: B00801/70265 (JRV/JAV)
; CURRENT APPLICATION NUMBER: US/10/236,031B
; CURRENT FILING DATE: 2002-09-05
; PRIOR APPLICATION NUMBER: US 60/317,389
; PRIOR FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: US 60/407,431
; PRIOR FILING DATE: 2002-08-30
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 69
; LENGTH: 7680
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-236-031B-69

Query Match          19.1%; Score 316.8; DB 6; Length 7680;
Best Local Similarity 99.4%; Pred. No. 4.2e-79;
Matches 318; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1339 GTCAGGCGCAGCAAAATGGTTTCAGCCCGCAGTCCCGGTGCTGTCAAGCAAGCCCG 1398
Db 12 GGCAGGCTCAGCAAAATGGTTTCAGCCCGCAGTCCCGGTGCTGTCAAGCAAGCCCG 71

Qy 1399 GTTGTATGACAAATGGAAACACTATCAGATAAATCAACAGTGGGAGCGGACCTACCTAG 1458
Db 72 GTTGTATGACAAATGGAAACACTATCAGATAAATCAACAGTGGGAGCGGACCTACCTAG 131
```

```

RESULT 15
US-10-182-936A-75
; Sequence 75, Application US/10182936A
; Publication No. US20040038860A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Kristina M.
; APPLICANT: Anisowicz, Anthony
; APPLICANT: Bhat, Bheem
; APPLICANT: Damagnez, Veronique
; APPLICANT: Robinson, John
; APPLICANT: Yaworsky, Paul
; TITLE OF INVENTION: Reagents and Method for Modulating DKK-Mediated Interactions
; FILE REFERENCE: 032796-143
; CURRENT APPLICATION NUMBER: US/10/182,936A
; CURRENT FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: PCT/US02/15982
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: US 60/291,311
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: US 60/353,058
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: US 60/361,293
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 216
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 75
; LENGTH: 7680
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-182-936A-75

Query Match      19.1%; Score 316.8; DB 7; Length 7680;
Best Local Similarity 99.4%; Pred. No. 4.2e-79;
Matches 318; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy      1339  GTCAGCGCAGCAATGGTTTCAGCCCCAGTCCCGGTGGCTGTCAAGCAAGCCCG 1398
Db      12   GCAGCGTCAGCAATGGTTTCAGCCCCAGTCCCGGTGGCTGTCAAGCAAGCCCG 71

Qy      1399  GTTGTATTGACAATGGAAACACTATCAGATAAATCAACAGTGGGAGCGGACCTACCTAG 1458
Db      72   GTTGTATTGACAATGGAAACACTATCAGATAAATCAACAGTGGGAGCGGACCTACCTAG 131

Qy      1459  GTAATGTGTGTTTGTACTTGTATTAGGAGGAAGCCGAGGTTTTAACTGCGAAGTAAC 1518
Db      132   GTAAATGTGTGTTTGTACTTGTATTAGGAGGAAGCCGAGGTTTTAACTGCGAAGTAAC 191

Qy      1519  CTGAGAGCTGAAGAGACTGCTTTTGACAACTACACTGGGACACTTACCGAGTGGTGACA 1578
Db      192   CTGAAGAGCTGAAGAGACTTGTCTTTGACAACTACACTGGGACACTTACCGAGTGGTGACA 251

Qy      1579  CTTATGAGCGTCCTAAAGACTCCATGATCTGGGACTGTACTTCATCTGGGCTGGGCGAG 1638
Db      252   CTTATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACTTCATCTGGGCTGGGCGAG 311

Qy      1639  GGAGAAATGAAGCTGTACCATC 1658
Db      312   GGAGAAATGAAGCTGTACCATC 331

```

Search completed: February 1, 2006, 14:21:19
Job time : 1373.44 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2006, 18:17:59 ; Search time 307.097 Seconds
(without alignments)
4492.841 Million cell updates/sec

Title: US-09-940-235-10
Perfect score: 1661
Sequence: 1 gcaacccgcagcctagcc.....gaataagctgaccatctaa 1661

Scoring table: IDENTITY NUC
Gap 10.0, Gapext 1.0
Searched: 6059551 seqs, 415333918 residues
Total number of hits satisfying chosen parameters: 12119102

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA New:
1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
4: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
6: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
7: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
8: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:
9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:
10: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:
11: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1661	100.0	1661	7	US-10-631-558-10
2	1282.2	77.2	1377	7	US-10-631-558-5
3	1219.4	73.4	1327	7	US-10-631-558-6
4	1185	71.3	1541	7	US-10-631-558-9
5	1165.8	70.2	2096	7	US-10-631-558-12
6	1150.2	69.2	1782	7	US-10-631-558-11
7	1148.6	69.2	1245	7	US-10-631-558-1
8	316.4	19.0	777	7	US-10-631-558-3
9	312	18.8	8232	7	US-10-821-234-693
10	311.6	18.8	2443	7	US-10-995-561-114
11	311.6	18.8	2488	7	US-10-995-561-105
12	311.6	18.8	6510	7	US-10-995-561-112
13	311.6	18.8	7823	7	US-10-995-561-117
14	311.6	18.8	7848	7	US-10-995-561-111
15	311.6	18.8	7935	7	US-10-995-561-113
16	311.6	18.8	7959	7	US-10-995-561-108
17	311.6	18.8	8013	7	US-10-995-561-104
18	311.6	18.8	8155	7	US-10-995-561-116
19	311.6	18.8	8226	7	US-10-995-561-107
20	311.6	18.8	8278	7	US-10-995-561-106
21	311.6	18.8	8332	7	US-10-995-561-110
22	311.6	18.8	8371	7	US-10-995-561-109

23	255	15.4	8404	8	US-11-136-527-2446	Sequence 2446, Ap
24	135.8	8.2	87672	7	US-10-995-561-13237	Sequence 13237, A
25	133.6	8.0	201	7	US-10-995-561-4446	Sequence 4446, Ap
26	133.6	8.0	201	7	US-10-995-561-4450	Sequence 4450, Ap
27	133.6	8.0	201	7	US-10-995-561-4468	Sequence 4468, Ap
28	133.6	8.0	201	7	US-10-995-561-4486	Sequence 4486, Ap
29	133.6	8.0	201	7	US-10-995-561-4503	Sequence 4503, Ap
30	133.6	8.0	201	7	US-10-995-561-4522	Sequence 4522, Ap
31	133.6	8.0	201	7	US-10-995-561-4541	Sequence 4541, Ap
32	133.6	8.0	201	7	US-10-995-561-4560	Sequence 4560, Ap
33	133.6	8.0	201	7	US-10-995-561-4575	Sequence 4575, Ap
34	133.6	8.0	201	7	US-10-995-561-4592	Sequence 4592, Ap
35	133.6	8.0	201	7	US-10-995-561-4596	Sequence 4596, Ap
36	133.6	8.0	201	7	US-10-995-561-4621	Sequence 4621, Ap
37	133.6	8.0	201	7	US-10-995-561-4640	Sequence 4640, Ap
38	133.6	8.0	201	7	US-10-995-561-27119	Sequence 27119, A
39	99	6.0	4974	8	US-11-094-586-17	Sequence 17, Appl
40	99	6.0	4974	8	US-11-076-733-86	Sequence 86, Appl
41	99	6.0	8966	8	US-11-076-733-87	Sequence 87, Appl
42	98	5.9	227	8	US-11-137-395-9	Sequence 9, Appli
43	98	5.9	5303	8	US-11-149-403-34	Sequence 34, Appl
44	98	5.9	7085	8	US-11-149-403-33	Sequence 33, Appl
45	98	5.9	8031	8	US-11-149-403-35	Sequence 35, Appl

ALIGNMENTS

RESULT 1
US-10-631-558-10
; Sequence 10, Application US/10631558
; Publication No. US20050260598A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammara
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/10/631.558
; PRIOR FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US/09/940.235
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 1661
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hybrid cassette
US-10-631-558-10

Query Match 100.0%; Score 1661; DB 7; Length 1661;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1661; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCAACCCGCGCCTAGCCGGTCTTCAACGAGGAGCAGATCATGCGACCCGTTGG 60
Db 1 GCAACCCGCGCCTAGCCGGTCTTCAACGAGGAGCAGATCATGCGACCCGTTGG 60
Qy 61 CCAGGACCCACCGTCCCGGATCTCGATCCCGGAATTAATACGACTCACTATAGG 120

Db 61 CCAGGACCAACGCTGCCGAGATCTCGATCCCGGAAATTAATACGACTCACTATAGG 120
 Qy 121 AGACCAACAGGTTTCCCTCTAGAAATAATTTTGTATTTAACTTTAAGAAGGAGATATACCA 180
 Db 121 AGACCAACAGGTTTCCCTCTAGAAATAATTTTGTATTTAACTTTAAGAAGGAGATATACCA 180
 Qy 181 TGAATGCTGGACTGAGTGGCTGCTAGAGCGTCCATCTGTCAACACAGCCAAATTTGGTTG 240
 Db 181 TGAATGCTGGACTGAGTGGCTGCTAGAGCGTCCATCTGTCAACACAGCCAAATTTGGTTG 240
 Qy 241 TTAGCGTTGCTGGTACTGTTGAGGGAAGCAATCAAGACATTAGTCTTAAATTTTTGAAA 300
 Db 241 TTAGCGTTGCTGGTACTGTTGAGGGAAGCAATCAAGACATTAGTCTTAAATTTTTGAAA 300
 Qy 301 TCGATCTAATCAACAGCTGCTCATGGAGGAAGACAGAGCGCTTAAGTCCAAAT 360
 Db 301 TCGATCTAATCAACAGCTGCTCATGGAGGAAGACAGAGCGCTTAAGTCCAAAT 360
 Qy 361 CAAAACCAATTTGCTACTGATAGTGGCGCATGTGCATATAAATTTGAGAAAGCTGACTTAC 420
 Db 361 CAAAACCAATTTGCTACTGATAGTGGCGCATGTGCATATAAATTTGAGAAAGCTGACTTAC 420
 Qy 421 TAAAGGCTATTCAAGAACTTGAATGATCGTAAACGTCCACAGTAAACGACTACTTTGAGG 480
 Db 421 TAAAGGCTATTCAAGAACTTGAATGATCGTAAACGTCCACAGTAAACGACTACTTTGAGG 480
 Qy 481 TCATTGATTTTCAAGCGATGCAACCAATTAATGATCGAATCGAAGCGAGGCTACTTTGCTG 540
 Db 481 TCATTGATTTTCAAGCGATGCAACCAATTAATGATCGAATCGAAGCGAGGCTACTTTGCTG 540
 Qy 541 ACAAGATGTTTGGTAACTTGGCGACCACTGTCCAAAGAAATTTTCTTAAGCGGAC 600
 Db 541 ACAAGATGTTTGGTAACTTGGCGACCACTGTCCAAAGAAATTTTCTTAAGCGGAC 600
 Qy 601 ATGTGCGGTTAGACATATAAGAAACCAATACAAACCAAGCGAAATCTGTTGATG 660
 Db 601 ATGTGCGGTTAGACATATAAGAAACCAATACAAACCAAGCGAAATCTGTTGATG 660
 Qy 661 TGGAAATATCTGATCAAGTTACTTCCCTTAAACCTGATGACGATTTTCAAGCAGGCTCA 720
 Db 661 TGGAAATATCTGATCAAGTTACTTCCCTTAAACCTGATGACGATTTTCAAGCAGGCTCA 720
 Qy 721 AAGTACTAAGCTATTGAAAAACATAGCTATCGTGCACACCATCTCAAGAAATTC 780
 Db 721 AAGTACTAAGCTATTGAAAAACATAGCTATCGTGCACACCATCTCAAGAAATTC 780
 Qy 781 TAGCTCAAGCAAAAGCAATTTTAAACAAACCAAGCGCTATAGGATTTATGAACGTG 840
 Db 781 TAGCTCAAGCAAAAGCAATTTTAAACAAACCAAGCGCTATAGGATTTATGAACGTG 840
 Qy 841 ACTCCTCAATCTGATCAATGACAAATGACATTTTCCGTAGCATTTTCAATGGATCAAG 900
 Db 841 ACTCCTCAATCTGATCAATGACAAATGACATTTTCCGTAGCATTTTCAATGGATCAAG 900
 Qy 901 AGTTTACTTACCGTGTAAATATCGGAAACAGCTTATAGGATCAATTAATTAATCTGGTC 960
 Db 901 AGTTTACTTACCGTGTAAATATCGGAAACAGCTTATAGGATCAATTAATTAATCTGGTC 960
 Qy 961 TGAATGAAGAAATAAACACACTGATCTGATGAAATATAGCTCTTAAATAAAG 1020
 Db 961 TGAATGAAGAAATAAACACACTGATCTGATGAAATATAGCTCTTAAATAAAG 1020
 Qy 1021 GGGAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTTCAACATCAATACG 1080
 Db 1021 GGGAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTTCAACATCAATACG 1080
 Qy 1081 TTGATGTCGATACCAAGAAATGCTTAAAGAGTGAAGGAGCTCTTAAAGCTAGGAACTA 1140
 Db 1081 TTGATGTCGATACCAAGAAATGCTTAAAGAGTGAAGGAGCTCTTAAAGCTAGGAACTA 1140
 Qy 1141 ACTTAGACTTCAAGATTTATAGATCTCTGATTAAGGCTTAACTACTCTACAAATC 1200
 Db 1141 ACTTAGACTTCAAGATTTATAGATCTCTGATTAAGGCTTAACTACTCTACAAATC 1200

Qy 1201 TCGATGCTTTTGGTATTATGGAATATACCTTAACCTGAAAAAGTAGAGATTAATCAGCATG 1260
 Db 1201 TCGATGCTTTTGGTATTATGGAATATACCTTAACCTGAAAAAGTAGAGATTAATCAGCATG 1260
 Qy 1261 ACACCAACCGTATCATAAACCGTTTATATGGCAAGCGACCCGAAAGGAGAGAAATCTAGCT 1320
 Db 1261 ACACCAACCGTATCATAAACCGTTTATATGGCAAGCGACCCGAAAGGAGAGAAATCTAGCT 1320
 Qy 1321 ATCATTTAGCCGTTGGTTCAGCGCAGCAAAATGGTTTCAGCCCCAGTCCCGGTGGCTG 1380
 Db 1321 ATCATTTAGCCGTTGGTTCAGCGCAGCAAAATGGTTTCAGCCCCAGTCCCGGTGGCTG 1380
 Qy 1381 TCAGTCAAAAGCAAGCCCGGTTTGTATGCAAAATGAAAAACATATCAGATAAATCAACAGT 1440
 Db 1381 TCAGTCAAAAGCAAGCCCGGTTTGTATGCAAAATGAAAAACATATCAGATAAATCAACAGT 1440
 Qy 1441 GGGAGCGGACTACTAGGTAATGTGTGGTTTGTACTTGTATGGAGGAAGCCGAGGTT 1500
 Db 1441 GGGAGCGGACTACTAGGTAATGTGTGGTTTGTACTTGTATGGAGGAAGCCGAGGTT 1500
 Qy 1501 TTAACCTGCAAAAGTAACCTGAAAGCTGAAGAGCTTGGTTTGAACAAGTACACTGGGAACA 1560
 Db 1501 TTAACCTGCAAAAGTAACCTGAAAGCTGAAGAGCTTGGTTTGAACAAGTACACTGGGAACA 1560
 Qy 1561 CTTTACCGAGTGGGTGACACTTATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACCT 1620
 Db 1561 CTTTACCGAGTGGGTGACACTTATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACCT 1620
 Qy 1621 GCATCGGGGCTGGCGAGGAGGAATAAGCTGTACCATCTAA 1661
 Db 1621 GCATCGGGGCTGGCGAGGAGGAATAAGCTGTACCATCTAA 1661

RESULT 2
 US-10-631-558-5
 ; Sequence 5, Application US/10631558
 ; Publication No. US20050260598A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kumar, Rajesh
 ; APPLICANT: Sahni, Girish
 ; APPLICANT: Roy, Chait
 ; APPLICANT: Rajagopal, Kammar
 ; APPLICANT: Nihalani, Deepak
 ; APPLICANT: Sundaram, Vasudha
 ; APPLICANT: Yadav, Mahavir
 ; TITLE OF INVENTION: NOVEL CILOT-SPECIFIC STREPTOKINASE
 ; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
 ; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
 ; TITLE OF INVENTION: PROTEIN
 ; FILE REFERENCE: 07064-009002
 ; CURRENT APPLICATION NUMBER: US/10/631,558
 ; CURRENT FILING DATE: 2003-07-31
 ; PRIOR APPLICATION NUMBER: US/09/940,235
 ; PRIOR FILING DATE: 2002-04-09
 ; PRIOR APPLICATION NUMBER: 09/471,349
 ; PRIOR FILING DATE: 1999-12-23
 ; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
 ; PRIOR FILING DATE: 1998-12-24
 ; NUMBER OF SEQ ID NOS: 28
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 5
 ; LENGTH: 1377
 ; TYPE: DNA
 ; ORGANISM: Streptococcus equisimilis
 US-10-631-558-5

Query Match 77.2%; Score 1282.2; DB 7; Length 1377;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1284; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Qy 51 GCACCGTGGCCAGGACCCCAACGCTGCCGAGATCTCGATCCCGGAAATTAATCAGCAT 110

Db 1 GCACCCGTCGACGACCCACGCTGCCGAGATCTCGATCCCGGAAATTAATACGACT 60
 QY 111 CACTATAGGAGACCAACGCGTTTCCCTCTAGAAATAATTTGTTAACTTTAAGAAGG 170
 Db 61 CACTATAGGAGACCAACGCGTTTCCCTCTAGAAATAATTTGTTAACTTTAAGAAGG 120
 QY 171 AGATATACCATGATTGCTGACCTGAGTGGCTGTAGACCGTCATCTGTCAACAACAGC 230
 Db 121 AGATATACCATGATTGCTGACCTGAGTGGCTGTAGACCGTCATCTGTCAACAACAGC 180
 QY 231 CAATTGGTCTGTAGCGTTGCTGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAA 290
 Db 181 CAATTGGTCTGTAGCGTTGCTGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAA 240
 QY 291 TTTTGTGAATCGATCTAACATACGACCTGCTCATGGAGGAAAGACAGACCAAGGCTTA 350
 Db 241 TTTTGTGAATCGATCTAACATACGACCTGCTCATGGAGGAAAGACAGACCAAGGCTTA 300
 QY 351 AGTCCAAATCAAAACCAATTTGCTACTGTAGTGGCGGATGTCAATATACTTTAGAAA 410
 Db 301 AGTCCAAATCAAAACCAATTTGCTACTGTAGTGGCGGATGTCAATATACTTTAGAAA 360
 QY 411 GCTGACTTAAAGGCTATTCAAGACCAATTTGATCGCTAAACGCTTAAACGACGAC 470
 Db 361 GCTGACTTAAAGGCTATTCAAGACCAATTTGATCGCTAAACGCTTAAACGACGAC 420
 QY 471 TACTTTGAGGTCAATGATTTTGAAGCGATGCAACCAATTAATCTGATCGAAACGCAAGTC 530
 Db 421 TACTTTGAGGTCAATGATTTTGAAGCGATGCAACCAATTAATCTGATCGAAACGCAAGTC 480
 QY 531 TACTTTGCTGCAAAAGATGTTTCGGTAACTTGCAGCCCAACCTGCTCCAGAAATTTTG 590
 Db 481 TACTTTGCTGCAAAAGATGTTTCGGTAACTTGCAGCCCAACCTGCTCCAGAAATTTTG 540
 QY 591 CTAAGCGGACATGTCGCGTTAGACCATATAAGAAAAACAATACAAAACCAAGCGAAA 650
 Db 541 CTAAGCGGACATGTCGCGTTAGACCATATAAGAAAAACAATACAAAACCAAGCGAAA 600
 QY 651 TCTGTTGATGTGAATATACGTACAGTTTACTCCCTTAAACCTGATGAGTTTCAGA 710
 Db 601 TCTGTTGATGTGAATATACGTACAGTTTACTCCCTTAAACCTGATGAGTTTCAGA 660
 QY 711 CCAGGCTCAAGATATACGTATTTGAAACACCTAGCTATCGGTGACACCATCACATCT 770
 Db 661 CCAGGCTCAAGATATACGTATTTGAAACACCTAGCTATCGGTGACACCATCACATCT 720
 QY 771 CAAGAAATCTAGCTCAAGCAAAAGCATTTTAAACCAAAACCAACCCAGGCTATACGATT 830
 Db 721 CAAGAAATCTAGCTCAAGCAAAAGCATTTTAAACCAAAACCAACCCAGGCTATACGATT 780
 QY 831 TATGACGTCATCTCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTACCA 890
 Db 781 TATGACGTCATCTCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTACCA 840
 QY 891 ATGGATCAAGAGTTTACTTACCGTGTAAATAATCGGAAACCAAGCTTATAGGATCAATAA 950
 Db 841 ATGGATCAAGAGTTTACTTACCGTGTAAATAATCGGAAACCAAGCTTATAGGATCAATAA 900
 QY 951 AAATCTGGTCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTC 1010
 Db 901 AAATCTGGTCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTC 960
 QY 1011 CTTAAAAAGGGGAAAAGCGGTATGATCCCTTTGATGCGAGTCACTTTGAAACTGTTCAAC 1070
 Db 961 CTTAAAAAGGGGAAAAGCGGTATGATCCCTTTGATGCGAGTCACTTTGAAACTGTTCAAC 1020
 QY 1071 ATCAATAGCTTGTGATCGATACCAAGCAATTTGATGAGTGGAGCGCTTTAAGACT 1130
 Db 1021 ATCAATAGCTTGTGATCGATACCAAGCAATTTGATGAGTGGAGCGCTTTAAGACT 1080
 QY 1131 AGCGAACGTAACCTTAGACTTCAGAGATTTATAGATCTCTCGTATAGGCTTAACTACTC 1190
 Db 1081 AGCGAACGTAACCTTAGACTTCAGAGATTTATAGATCTCTCGTATAGGCTTAACTACTC 1140

QY 1191 TACAACAATCTCGATCTTTGGTATTATGGAATATACCTTAATCTGGAAGAGTAGAGAT 1250
 Db 1141 TACAACAATCTCGATCTTTGGTATTATGGAATATACCTTAATCTGGAAGAGTAGAGAT 1200
 QY 1251 AATCAGATGACCAACCGGTATCAACCGTTTATATGGCAACGACCCGAAGAGAG 1310
 Db 1201 AATCAGATGACCAACCGGTATCAACCGTTTATATGGCAACGACCCGAAGAGAG 1260
 QY 1311 AATGTAGCTATCATTTAGCCGGTGGT 1337
 Db 1261 AATGTAGCTATCATTTAGCCCTATGAT 1287

RESULT 3

US-10-631-558-6
 ; Sequence 6, Application US/10631558
 ; Publication No. US20050260598A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kumar, Rajesh
 ; APPLICANT: Sabni, Girish
 ; APPLICANT: Roy, Chait
 ; APPLICANT: Rajagopal, Kammar
 ; APPLICANT: Nihalani, Deepak
 ; APPLICANT: Sundaram, Vasudha
 ; APPLICANT: Yadav, Mahavir
 ; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
 ; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
 ; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
 ; TITLE OF INVENTION: PROTEIN
 ; FILE REFERENCE: 07064-009002
 ; CURRENT APPLICATION NUMBER: US/10/631,558
 ; PRIOR FILING DATE: 2003-07-31
 ; PRIOR APPLICATION NUMBER: US/09/940,235
 ; PRIOR FILING DATE: 2002-04-09
 ; PRIOR APPLICATION NUMBER: 09/471,349
 ; PRIOR FILING DATE: 1999-12-23
 ; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
 ; PRIOR FILING DATE: 1998-12-24
 ; NUMBER OF SEQ ID NOS: 28
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 6
 ; LENGTH: 1327
 ; TYPE: DNA
 ; ORGANISM: Streptococcus equisimilis
 US-10-631-558-6

Query Match 73.4%; Score 1219.4; DB 7; Length 1327;
 Best Local Similarity 99.1%; Pred. No. 0;
 Matches 1226; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
 QY 101 TAATACGACTCACTATAGGAGAGACCAACGCGTTTCCCTCTAGAAATAATTTGTTTAA 160
 Db 1 TAATACGACTCACTATAGGAGAGACCAACGCGTTTCCCTCTAGAAATAATTTGTTTAA 60
 QY 161 TTTAAGAGGAGATATACCATGATGCTGGACCTGAGTGGCTGTAGACCGTCATCTGT 220
 Db 61 TTTAAGAGGAGATATACCATGATGCTGGACCTGAGTGGCTGTAGACCGTCATCTGT 120
 QY 221 CAACAACAGCCAAATTTGTTAGCTTGTAGCTGCTGATGCTGCTCATGAGGAGAAAGACA 280
 Db 121 AAATAACAGCCAAATTTGTTAGCTTGTAGCTGCTGATGCTGCTCATGAGGAGAAAGACA 180
 QY 281 TAGTCTTAAATTTTGAATTCGATCAACATCAACGCTCTCATGAGGAGAAAGACA 340
 Db 181 TAGTCTTAAATTTTGAATTCGATCAACATCAACGCTCTCATGAGGAGAAAGACA 240
 QY 341 GCAAGGCTTAAGTCCAAATCAAAACCAATTTGCTGATGATGTCGCGATGTCACTAA 400
 Db 241 GCAAGGCTTAAGTCCAAATCAAAACCAATTTGCTGATGATGTCGCGATGTCACTAA 300
 QY 401 ACTTGAGAGGCTGACTTACTTAAAGGCTTATCAAGAACAAATTTGATCGTAACTCCAC 460

```

Db 301 ACTTGAGAAAGTGACTTAAAGGCTATTCAGAAACAATTGATCGTCAACGTCACAG 360
Qy 461 TAAACACGACTACTTTGAGGTCAATGATTTTGAAGCGATGCAACCAATTAAGTCAAAA 520
Db 361 TAAACACGACTACTTTGAGGTCAATGATTTTGAAGCGATGCAACCAATTAAGTCAAAA 420
Qy 521 CGGCAAGGFTACTTTGCTGACAAAGATGGTTCGGTAAACCTTCCGACCAACCTGTCCA 580
Db 421 CGGCAAGGFTACTTTGCTGACAAAGATGGTTCGGTAAACCTTCCGACCAACCTGTCCA 480
Qy 581 AGAATTTTCTTAAGCGGACATGTCGGCTTAGACCAATATAAGAAACCAATACAAAA 640
Db 481 AGAATTTTCTTAAGCGGACATGTCGGCTTAGACCAATATAAGAAACCAATACAAAA 540
Qy 641 CCAAGCGAATCTGTTGATGTGAATATACGTACACAGTTTACCTCCCTTAAACCTGTATGA 700
Db 541 CCAAGCGAATCTGTTGATGTGAATATACGTACACAGTTTACCTCCCTTAAACCTGTATGA 600
Qy 701 CGATTTTCAGACCAAGTCTCAAGATATTAAGCTATTTGAAAAACACTAGCTATCGGTGACAC 760
Db 601 CGATTTTCAGACCAAGTCTCAAGATATTAAGCTATTTGAAAAACACTAGCTATCGGTGACAC 660
Qy 761 CATCATCTCAAGAAATTAAGTCTCAAGCAAAAGCATTTTAAACCAAAACCAACCCAGG 820
Db 661 CATCATCTCAAGAAATTAAGTCTCAAGCAAAAGCATTTTAAACCAAAACCAACCCAGG 720
Qy 821 CTATACGATTTATGAAGGTGACTCTCAATCGTCACTCATGACAAATGACATTTTCCGTAC 880
Db 721 CTATACGATTTATGAAGGTGACTCTCAATCGTCACTCATGACAAATGACATTTTCCGTAC 780
Qy 881 GATTTTACCAATCGGATCAAGAGTTTACTTACCGTGTAAAAATCGGGAACCAAGCTTATAG 940
Db 781 GATTTTACCAATCGGATCAAGAGTTTACTTACCGTGTAAAAATCGGGAACCAAGCTTATAG 840
Qy 941 GATCAATAAAAAATCTGGTCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAA 1000
Db 841 GATCAATAAAAAATCTGGTCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAA 900
Qy 1001 ATATTACGCTTTAAAAAGGGGAAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAA 1060
Db 901 ATATTACGCTTTAAAAAGGGGAAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAA 960
Qy 1061 ACTGTTTACCATCAATACGTTGATGTCGATACCAAGCAATTCCTTAAAAAGTGCAGAGCT 1120
Db 961 ACTGTTTACCATCAATACGTTGATGTCGATACCAAGCAATTCCTTAAAAAGTGCAGAGCT 1020
Qy 1121 CTTAAACAGTAGCGAACGTAACCTTAGACTTTCAGAGATTTTATACGATCCTCGTGAAGGC 1180
Db 1021 CTTAAACAGTAGCGAACGTAACCTTAGACTTTCAGAGATTTTATACGATCCTCGTGAAGGC 1080
Qy 1181 TAAACTACTTCAACAATCTCGATGCTTTGGTATTTATGGAATATACCTTAACTGGA 1240
Db 1081 TAAACTACTTCAACAATCTCGATGCTTTGGTATTTATGGAATATACCTTAACTGGA 1140
Qy 1241 AGTAGAGGATATACGATGACCAACCGTATCATACCGTTTATATGGCGAGCGACC 1300
Db 1141 AGTAGAGGATATACGATGACCAACCGTATCATACCGTTTATATGGCGAGCGACC 1200
Qy 1301 CGAAGGAGAGATGCTAGCTATCATTTAGCCGGTGGT 1337
Db 1201 CGAAGGAGAGATGCTAGCTATCATTTAGCCCTATGAT 1237

```

RESULT 4

US-10-631-558-9

; Sequence 9, Application US/10631558

; Publication No. US20050260598A1

; GENERAL INFORMATION:

; APPLICANT: Kumar, Rajesh

; APPLICANT: Sahni, Girish

; APPLICANT: Roy, Chait

; APPLICANT: Rajagopal, Kammar

; APPLICANT: Nihalani, Deepak

```

; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/10/631,558
; CURRENT FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US/09/940,235
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1541
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hybrid cassette
US-10-631-558-9

```

Query Match 71.3%; Score 1185; DB 7; Length 1541;

Best Local Similarity 98.8%; Pred. No. 0; Mismatches 15; Indels 0; Gaps 0; Matches 1194; Conservative 0;

```

Qy 151 TTTGTTTAACTTTAAGAGGAGATATACCATGATTGCTGGACCTGAGTGGCTGCTAGACC 210
Db 1 TTTGTTTAACTTTAAGAGGAGATATACCATGATTGCTGGACCTGAGTGGCTGCTAGATC 60
Qy 211 GTCCATCTGTCAACAAGCCAAATGTTGTTAGCGTTGCTGGTACTGTTGAGGGAGCA 270
Db 61 GTCCCTCTGTAATAAACAAGCCAAATGTTGTTAGCGTTGCTGGTACTGTTGAGGGAGCA 120
Qy 271 ATCAAGACATTAAGTCTTAATTTTGAATCGATCTAAACATCAAGACCTGCTCATGGAG 330
Db 121 ATCAAGACATTAAGTCTTAATTTTGAATCGATCTAAACATCAAGACCTGCTCATGGAG 180
Qy 331 GAAAGACAGACGAGGCTTAAGTCCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCA 390
Db 181 GAAAGACAGACGAGGCTTAAGTCCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCA 240
Qy 391 TGTCAACATAAATTTGAGAAAGCTGACTTAATAAGGCTATTCAAGAGCTATTCAAGAAACA 450
Db 241 TGTCAACATAAATTTGAGAAAGCTGACTTAATAAGGCTATTCAAGAGCTATTCAAGAAACA 300
Qy 451 ACCTCCACAGTACGACGACTACTTTGAGGTCATTTGAGGTCATTTGAGGTCATTTGAGGTC 510
Db 301 ACCTCCACAGTACGACGACTACTTTGAGGTCATTTGAGGTCATTTGAGGTCATTTGAGGTC 360
Qy 511 CTGATCGAAACGGCAAGGCTACTTTGCTGCAAAAGATGGTTCGGTAACTTTCGCGACCC 570
Db 361 CTGATCGAAACGGCAAGGCTACTTTGCTGCAAAAGATGGTTCGGTAACTTTCGCGACCC 420
Qy 571 AACCTGTCCAAAGATTTTGTCTAAGCGCAATGTGCGGCTTAGACCATATAAAGAAAAAC 630
Db 421 AACCTGTCCAAAGATTTTGTCTAAGCGCAATGTGCGGCTTAGACCATATAAAGAAAAAC 480
Qy 631 CAATACAAAACCAAGCGAATCTGTTGATGTGGAATATCTGTTACAGTACTTCTCCCTTAA 690
Db 481 CAATACAAAACCAAGCGAATCTGTTGATGTGGAATATCTGTTACAGTACTTCTCCCTTAA 540
Qy 691 ACCTGATGACGATTTTCAGACCGAGTCTCAAAAGATCTAAAGCTATTGAAACACTTAGCTA 750
Db 541 ACCTGATGACGATTTTCAGACCGAGTCTCAAAAGATCTAAAGCTATTGAAACACTTAGCTA 600
Qy 751 TCGGTGACACCATCAATCTCAAGAAATTAAGTCTCAAGCAACAAGCAATTTTAAACAAAA 810
Db 601 TCGGTGACACCATCAATCTCAAGAAATTAAGTCTCAAGCAACAAGCAATTTTAAACAAAA 660

```

```

QY 811 ACCACCCAGCTATACGATTATGAAGTGAATCTCTCAATCGTCACTCATGACAATGACA 870
Db 661 ACCACCCAGCTATACGATTATGAAGTGAATCTCTCAATCGTCACTCATGACAATGACA 720
QY 871 TTTTCCGTACGATTATACCAATGATCAAGATTTTACTTACCGGTAAATAATCGGAC 930
Db 721 TTTTCCGTACGATTATACCAATGATCAAGATTTTACTTACCGGTAAATAATCGGAC 780
QY 931 AAGCTTATAGGATCAATAAAAAATCTGGTCTGAATGAAGAAATAAAACAACACTGACCTGA 990
Db 781 AAGCTTATAGGATCAATAAAAAATCTGGTCTGAATGAAGAAATAAAACAACACTGACCTGA 840
QY 991 TCTCTGAGAAATATAGCTCTTAAAAAAAGGGGAAAGCCGTATGATCCCTTTGATCGCA 1050
Db 841 TCTCTGAGAAATATAGCTCTTAAAAAAAGGGGAAAGCCGTATGATCCCTTTGATCGCA 900
QY 1051 GTCACCTGAAACGTTCACCATCAATACCTTGAATGTCGATACCAACGAATTTGCTAAAA 1110
Db 901 GTCACCTGAAACGTTCACCATCAATACCTTGAATGTCGATACCAACGAATTTGCTAAAA 960
QY 1111 GTGAGCAGCTCTTAAACAGCTAGCGAAGCTAACTTAGACTTCAGAGATTTATACGATCCTC 1170
Db 961 GTGAGCAGCTCTTAAACAGCTAGCGAAGCTAACTTAGACTTCAGAGATTTATACGATCCTC 1020
QY 1171 GTGATAGGCTAACTACTCTCAACAATCTCGATGCTTTTGGTATTTATGAGCTATACCT 1230
Db 1021 GTGATAGGCTAACTACTCTCAACAATCTCGATGCTTTTGGTATTTATGAGCTATACCT 1080
QY 1231 TAACTGAAAGTAGAGGATAATCAGATGACACCAACCGTATCATACCGTTTATATGG 1290
Db 1081 TAACTGAAAGTAGAGGATAATCAGATGACACCAACCGTATCATACCGTTTATATGG 1140
QY 1291 GCAAGCGACCGAAGGAGAGATGCTAGCTATCATTTAGCCGGTGGTGGTCAAGCGCAGC 1350
Db 1141 GCAAGCGACCGAAGGAGAGATGCTAGCTATCATTTAGCTGGTGGTGGCGGCGCAAC 1200
QY 1351 AAATGGTTC 1359
Db 1201 AGATTGTAC 1209

```

RESULT 5

```

; Sequence 12, Application US/10631558
; Publication No. US20050260598A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammar
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/10/631,558
; PRIOR FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US/09/940,235
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2096
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:

```

OTHER INFORMATION: Hybrid cassette
US-10-631-558-12

Query Match 70.2%; Score 1165.8; DB 7; Length 2096;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1170; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

QY 183 ATTGCTGACCTGAGTGGCTGTAGACCGTCTCATCTGTCAACAACAGCAATGTTGTT 242
Db 588 ATTGCTGACCTGAGTGGCTGTAGACCGTCTCATCTGTCAACAACAGCAATGTTGTT 647
QY 243 AGCGTTGCTGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTTAAATTTTGAATC 302
Db 648 AGCGTTGCTGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTTAAATTTTGAATC 707
QY 303 GATCTAACATCAACGACCTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAGTCCAAATCA 362
Db 708 GATCTAACATCAACGACCTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAGTCCAAATCA 767
QY 363 AAACCATTTGCTACTGTAGTGGCGGATGTACATAAATTTGAGAAAGCTGACTTACTA 422
Db 768 AAACCATTTGCTACTGTAGTGGCGGATGTACATAAATTTGAGAAAGCTGACTTACTA 827
QY 423 AAGGCTATTCAAGAACAAATTGATCGCTAAACGTCACAGTAACGACGACTACTTTGAGTC 482
Db 828 AAGGCTATTCAAGAACAAATTGATCGCTAAACGTCACAGTAACGACGACTACTTTGAGTC 887
QY 483 ATTGATTTTGAACGGATGCAACCATTTACTGATCGAAACGGCAAGGCTACTTTGCTGAC 542
Db 888 ATTGATTTTGAACGGATGCAACCATTTACTGATCGAAACGGCAAGGCTACTTTGCTGAC 947
QY 543 AAAGATGGTTCGGTAACTTCCCGACCAACCTGTCCCAAGAAATTTTGTCTAAGCGGACAT 602
Db 948 AAAGATGGTTCGGTAACTTCCCGACCAACCTGTCCCAAGAAATTTTGTCTAAGCGGACAT 1007
QY 603 GTGCGCGTTAGACCATATAAAGAAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 662
Db 1008 GTGCGCGTTAGACCATATAAAGAAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 1067
QY 663 GAATATATCTGTACAGTTTACTTCCCTTAAACCCCTGATGACGATTTTCAACAGGCTCAAA 722
Db 1068 GAATATATCTGTACAGTTTACTTCCCTTAAACCCCTGATGACGATTTTCAACAGGCTCAAA 1127
QY 723 GATACTAAGCTATTTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTACTA 782
Db 1128 GATACTAAGCTATTTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTACTA 1187
QY 783 GCTCAAGCACAAAGCATTTTAAACAAAAACCAACCCAGGCTATACGATTTATGAACGTGAC 842
Db 1188 GCTCAAGCACAAAGCATTTTAAACAAAAACCAACCCAGGCTATACGATTTATGAACGTGAC 1247
QY 843 TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 902
Db 1248 TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 1307
QY 903 TTTACTTACCGTGTAAAAAATCGGGAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 962
Db 1308 TTTACTTACCGTGTAAAAAATCGGGAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 1367
QY 963 AATGAAGAAATAAACACACTGCTCTCTGAGAAATATTTACGTCCTTAAAAAAGGG 1022
Db 1368 AATGAAGAAATAAACACACTGCTCTCTGAGAAATATTTACGTCCTTAAAAAAGGG 1427
QY 1023 GAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTTGAAACCTGTTTACCATCAAAATAGCT 1082
Db 1428 GAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTTGAAACCTGTTTACCATCAAAATAGCT 1487
QY 1083 GATGTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTTAAACAGCTAGCGAACGTAAC 1142
Db 1488 GATGTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTTAAACAGCTAGCGAACGTAAC 1547
QY 1143 TTAGACTTTCAGAGATTTTATACGATCTCTCGTGATAGGCTAAACTTACTCTACAAATCTC 1202

```

Db 1548 TTAGACTTCAGAGATTATATACGATCCTCGTGATAGGCTAAACTACTCTACAACAATCTC 1607
Qy 1203 GATGCTTTTGGTATTATGAGTATACCTTAACCTGGAAGCTAGAGGATAATACGATGAC 1262
Db 1608 GATGCTTTTGGTATTATGAGTATACCTTAACCTGGAAGCTAGAGGATAATACGATGAC 1667
Qy 1263 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAAATGCTAGCTAT 1322
Db 1668 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAAATGCTAGCTAC 1727
Qy 1323 CATTTAGCGGTGGTGTGAGCGCAGCAAAATGGTTC 1359
Db 1728 CATTTAGCGGTGGTGTGAGCGCAGCAAAATGGTAC 1764

RESULT 6

US-10-631-558-11
; Sequence 11, Application US/10631558
; Publication No. US20050260598A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammara
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/10/631,558
; CURRENT FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US/09/940,235
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 1782
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hybrid cassette
US-10-631-558-11

Query Match 69.2%; Score 1150.2; DB 7; Length 1782;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1152; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 183 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACGACCAATTTGGTGT 242
Db 538 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACGACCAATTTGGTGT 597
Qy 243 AGCGTGTGCTGCTACTCTGTAGGGGACCAATCAAGACATTAAGTCTTAAATTTTGAATC 302
Db 598 AGCGTGTGCTGCTACTCTGTAGGGGACCAATCAAGACATTAAGTCTTAAATTTTGAATC 657
Qy 303 GATCTAAATCATCAGACCTGCTATGAGGAGAAAGACAGACAGAGCGTCTTAAGTCAAAATCA 362
Db 658 GATCTAAATCATCAGACCTGCTATGAGGAGAAAGACAGACAGAGCGTCTTAAGTCAAAATCA 717
Qy 363 AAACCAATTTGCTACTCATAGTGGCGGATGTCAATAAATTCAGAAAGCTGACTACTA 422
Db 718 AAACCAATTTGCTACTCATAGTGGCGGATGTCAATAAATTCAGAAAGCTGACTACTA 777
Qy 423 AAGGCTATTCAAGAAACAATTTGATCGCTAACGTCACAGTAAACGACGACTACTTTGAGTTC 482
Db 778 AAGGCTATTCAAGAAACAATTTGATCGCTAACGTCACAGTAAACGACGACTACTTTGAGTTC 837

Qy 483 ATTGATTTTGCAGCGATGCAACCATTTACTGATCGAAACGGAAGGTCTACTTTGCTGAC 542
Db 838 ATTGATTTTGCAGCGATGCAACCATTTACTGATCGAAACGGAAGGTCTACTTTGCTGAC 897
Qy 543 AAAGATGGTTGCGTAAACCTTTGCCGACCCCAACCTCTGCCAAGAAATTTTGTCTAAGCGGACAT 602
Db 898 AAAGATGGTTGCGTAAACCTTTGCCGACCCCAACCTCTGCCAAGAAATTTTGTCTAAGCGGACAT 957
Qy 603 GTGGCGGTTAGACCATATAAAGAAAAACCAATACAAAACCAAGCGAATCTGTTGATGTG 662
Db 958 GTGGCGGTTAGACCATATAAAGAAAAACCAATACAAAACCAAGCGAATCTGTTGATGTG 1017
Qy 663 GAATATATCTGTACAGTTTACTTCCCTTAAACCCCTCATGACGATTTTTCAGACGAGTCTCAAA 722
Db 1018 GAATATATCTGTACAGTTTACTTCCCTTAAACCCCTCATGACGATTTTTCAGACGAGTCTCAAA 1077
Qy 723 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTAATA 782
Db 1078 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTAATA 1137
Qy 783 GCTCAAGCAACAAGCATTTTAAACAAAAACCCAGCGCTATACGATTTATGAAACGCTGAC 842
Db 1138 GCTCAAGCAACAAGCATTTTAAACAAAAACCCAGCGCTATACGATTTATGAAACGCTGAC 1197
Qy 843 TCCTCAATCTGCTCATGACAAATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 902
Db 1198 TCCTCAATCTGCTCATGACAAATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 1257
Qy 903 TTTACTTACCGTGTAAATAATCGGGAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 962
Db 1258 TTTACTTACCGTGTAAATAATCGGGAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 1317
Qy 963 AATGAAGAAATAAACAACATGATCTCTGAGAAATATTAAGTCTTTAAAAAAGGG 1022
Db 1318 AATGAAGAAATAAACAACATGATCTCTGAGAAATATTAAGTCTTTAAAAAAGGG 1377
Qy 1023 GAAAGCCGTATGATCCCTTTGATCGGAGTCACCTTTGAAACTGTTTCAACATCAATACGTT 1082
Db 1378 GAAAGCCGTATGATCCCTTTGATCGGAGTCACCTTTGAAACTGTTTCAACATCAATACGTT 1437
Qy 1083 GATGTCGATACCAACGAATTTCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAAACGTAAC 1142
Db 1438 GATGTCGATACCAACGAATTTCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAAACGTAAC 1497
Qy 1143 TTAGACTTCAGAGATTTTATACGATCTCTGATGAAGGCTAACTACTCTACAAATCTC 1202
Db 1498 TTAGACTTCAGAGATTTTATACGATCTCTGATGAAGGCTAACTACTCTACAAATCTC 1557
Qy 1203 GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAGTAGAGGATAATCACGATGAC 1262
Db 1558 GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAGTAGAGGATAATCACGATGAC 1617
Qy 1263 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAAATGCTAGCTAT 1322
Db 1618 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAAATGCTAGCTAT 1677
Qy 1323 CATTTAGCGGTGGT 1337
Db 1678 CATTTAGCGGTGGT 1692

RESULT 7

US-10-631-558-1
; Sequence 1, Application US/10631558
; Publication No. US20050260598A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammara
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha

APPLICANT: Yadav, Mahavir
TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
TITLE OF INVENTION: PROTEIN
FILE REFERENCE: 07064-009002
CURRENT APPLICATION NUMBER: US/10/631,558
PRIOR FILING DATE: 2003-07-31
PRIOR APPLICATION NUMBER: US/09/940,235
PRIOR FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: 09/471,349
PRIOR FILING DATE: 1999-12-23
PRIOR APPLICATION NUMBER: IN 3825/DEL/98
PRIOR FILING DATE: 1998-12-24
NUMBER OF SEQ ID NOS: 28
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 1245
TYPE: DNA
ORGANISM: Streptococcus equisimilis
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(1242)
US-10-631-558-1

Query Match : 69.2%; Score 1148.6; DB 7; Length 1245;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1151; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 183 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCCAATGGTGT 242
DB 1 ATTTGGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCCAATGGTGT 60
QY 243 ACGTTGCTGGTACTCTGAGGGGCGAATCAAGACATTAGTCTTAAATTTTGGAAATC 302
DB 61 ACGTTGCTGGTACTCTGAGGGGCGAATCAAGACATTAGTCTTAAATTTTGGAAATC 120
QY 303 GATCTAACTACGACCTGCTCATGGAGGAAAGACAGACGAAGCTTAAAGTCCAAATCA 362
DB 121 GATCTAACTACGACCTGCTCATGGAGGAAAGACAGACGAAGCTTAAAGTCCAAATCA 180
QY 363 AAACCAATTTGCTACTGATAGTGGCGGATGTCATATAAATTGAGAAAGCTGACTTACTA 422
DB 181 AAACCAATTTGCTACTGATAGTGGCGGATGTCATATAAATTGAGAAAGCTGACTTACTA 240
QY 423 AAGGCTATTCAAGAACAAATTGATCGCTAAGCTCAGAGTAACGACGACTACTTTGAGTGC 482
DB 241 AAGGCTATTCAAGAACAAATTGATCGCTAAGCTCAGAGTAACGACGACTACTTTGAGTGC 300
QY 483 ATTGATTTTGAAGCGATGCAACCATTTACTGATCGAAACGCAAGGTCTACTTTGCTGAC 542
DB 301 ATTGATTTTGAAGCGATGCAACCATTTACTGATCGAAACGCAAGGTCTACTTTGCTGAC 360
QY 543 AAAGATGGTTCGGTAACCTTTGCCGACCAACCTGTCCAGAATTTTGTCTAAGCGGACAT 602
DB 361 AAAGATGGTTCGGTAACCTTTGCCGACCAACCTGTCCAGAATTTTGTCTAAGCGGACAT 420
QY 603 GTGCGGTTAGACCATATAAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 662
DB 421 GTGCGGTTAGACCATATAAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 480
QY 663 GAATATAGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTCAA 722
DB 481 GAATATAGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTCAA 540
QY 723 GATCTAACTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTA 782
DB 541 GATCTAACTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTA 600
QY 783 GCTCAAGCAAAAGCAATTTTAAACAAAAACCCAGGCTATAGATTTATGAACGTGAC 842
DB 601 GCTCAAGCAAAAGCAATTTTAAACAAAAACCCAGGCTATAGATTTATGAACGTGAC 660

QY 843 TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 902
DB 661 TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 720
QY 903 TTTTACTTACCGTGTAAAAATCGGGAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 962
DB 721 TTTTACTTACCGTGTAAAAATCGGGAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 780
QY 963 AATGAAGAAATAAAACAACACATGATCTCTGAGAAATATTAGCTCTTAAAAAAGGG 1022
DB 781 AATGAAGAAATAAAACAACACATGATCTCTGAGAAATATTAGCTCTTAAAAAAGGG 840
QY 1023 GAAAAGCGGTATGATCCCTTTTGTGCGAGTCACTTTGAAACTGTTCAACATCAATCGTT 1082
DB 841 GAAAAGCGGTATGATCCCTTTTGTGCGAGTCACTTTGAAACTGTTCAACATCAATCGTT 900
QY 1083 GATGTCGATACCAAGCAATTTCTTAAAGTAGGACGCTTTAAACAGCTAGCGAAGTAAAC 1142
DB 901 GATGTCGATACCAAGCAATTTCTTAAAGTAGGACGCTTTAAACAGCTAGCGAAGTAAAC 960
QY 1143 TTAGACTTCAGAGATTTTATACGATCCTCGTATAGGCTTAACTCTTACAAACATCTC 1202
DB 961 TTAGACTTCAGAGATTTTATACGATCCTCGTATAGGCTTAACTCTTACAAACATCTC 1020
QY 1203 GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAGTAGAGGATAATCAGATGAC 1262
DB 1021 GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAGTAGAGGATAATCAGATGAC 1080
QY 1263 ACCAACCGTATCATAAACCGTTTATATGGGCAAGGACCCGGAAGGAGAGATGCTAGCTAT 1322
DB 1081 ACCAACCGTATCATAAACCGTTTATATGGGCAAGGACCCGGAAGGAGAGATGCTAGCTAT 1140
QY 1323 CATTTAGCCGGTGT 1337
DB 1141 CATTTAGCCGTATGAT 1155
RESULT 8
US-10-631-558-3
Sequence 3, Application US/10631558
Publication No. US20050260598A1
GENERAL INFORMATION:
APPLICANT: Kumar, Rajesh
APPLICANT: Sahni, Girish
APPLICANT: Roy, Chait
APPLICANT: Rajagopal, Kammar
APPLICANT: Nihalani, Deepak
APPLICANT: Sundaram, Vasudha
APPLICANT: Yadav, Mahavir
TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
TITLE OF INVENTION: PROTEIN
FILE REFERENCE: 07064-009002
CURRENT APPLICATION NUMBER: US/10/631,558
PRIOR FILING DATE: 2003-07-31
PRIOR APPLICATION NUMBER: US/09/940,235
PRIOR FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: 09/471,349
PRIOR FILING DATE: 1999-12-23
PRIOR APPLICATION NUMBER: IN 3825/DEL/98
PRIOR FILING DATE: 1998-12-24
NUMBER OF SEQ ID NOS: 28
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 777
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(777)
US-10-631-558-3

Query Match 19.0%; Score 316.4; DB 7; Length 777;
 Best Local Similarity 99.7%; Pred. No. 6.5e-89;
 Matches 317; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1341 CAGGCGCAGCAATGTTTCAGCCCGCAGTCCCGGTCGCTCAGTCAAGCAAGCCCGGT 1400
 DB 1 CAGGCTCAGCAATGTTTCAGCCCGCAGTCCCGGTCGCTCAGTCAAGCAAGCCCGGT 60

QY 1401 TGTATGACAAATGGAACCACTATCAGATAAATCAACAGTGGGAGCGGACCTACTAGGT 1460
 DB 61 TGTATGACAAATGGAACCACTATCAGATAAATCAACAGTGGGAGCGGACCTACTAGGT 120

QY 1461 AATGTTGTTGTTGTTACTTGTATGAGGAGCGGAGGTTTAACTCGGAAAGTAAACCT 1520
 DB 121 AATGTTGTTGTTGTTACTTGTATGAGGAGCGGAGGTTTAACTCGGAAAGTAAACCT 180

QY 1521 GAAGCTGAAGAGACTTGTCTTGAACAAGTACACTTGAAGAACTTACCGAGTGGGTGACACT 1580
 DB 181 GAAGCTGAAGAGACTTGTCTTGAACAAGTACACTTGAAGAACTTACCGAGTGGGTGACACT 240

QY 1581 TATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACCTGCGATCGGGGCTGGCGGAGG 1640
 DB 241 TATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACCTGCGATCGGGGCTGGCGGAGG 300

QY 1641 AGAATAAGCTGTACCATC 1658
 DB 301 AGAATAAGCTGTACCATC 318

RESULT 9
 US-10-821-234-693/c
 ; Sequence 693, Application US/10821234
 ; Publication No. US20050255114A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Labat, Ivan
 ; APPLICANT: Stache-Crain, Birgit
 ; APPLICANT: Andarmani, Susan
 ; APPLICANT: Tang, Y. Tom
 ; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
 ; FILE REFERENCE: 821A
 ; CURRENT APPLICATION NUMBER: US/10/821,234
 ; CURRENT FILING DATE: 2004-04-07
 ; PRIOR APPLICATION NUMBER: US 60/462,047
 ; PRIOR FILING DATE: 2003-04-07
 ; NUMBER OF SEQ ID NOS: 1704
 ; SOFTWARE: pc_seq_genes Version 1.0
 ; SEQ ID NO 693
 ; LENGTH: 8232
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-821-234-693

Query Match 18.8%; Score 312; DB 7; Length 8232;
 Best Local Similarity 98.4%; Pred. No. 6.7e-87;
 Matches 315; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1339 GTCAGGCGCAGCAATGTTTCAGCCCGCAGTCCCGGTCGCTCAGTCAAGCAAGCCCG 1398
 DB 7874 GGCAGGCTCAGCAATGTTTCAGCCCGCAGTCCCGGTCGCTCAGTCAAGCAAGCCCG 7815

QY 1399 GTTGTATGACAAATGGAACCACTATCAGATAAATCAACAGTGGGAGCGGACCTACCTAG 1458
 DB 7814 GTTGTATGACAAATGGAACCACTATCAGATAAATCAACAGTGGGAGCGGACCTACCTAG 7755

QY 1459 GTAATGTGTGTTGTTACTTGTATGAGGAGCGGAGGTTTAACTCGGAAAGTAAAC 1518
 DB 7754 GCAATGCGTGTGTTGTTACTTGTATGAGGAGCGGAGGTTTAACTCGGAGTAAAC 7695

QY 1519 CTGAAGCTGAAGAGACTTGTCTTGAACAAGTACACTGGGAACTTACCGAGTGGGTGACA 1578
 DB 7694 CTGAAGCTGAAGAGACTTGTCTTGAACAAGTACACTGGGAACTTACCGAGTGGGTGACA 7635

QY 1579 CTTATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACCTGCGGCTGGGCGAG 1638

Db 7634 CTTATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACCTGCGGCTGGGCGAG 7575

QY 1639 GGGAATAAGCTGTACCATC 1658
 Db 7574 GGGAATAAGCTGTACCATC 7555

RESULT 10
 US-10-995-561-114
 ; Sequence 114, Application US/10995561
 ; Publication No. US20050272054A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cargill, Michele et al.
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
 ; TITLE OF INVENTION: DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001559
 ; CURRENT APPLICATION NUMBER: US/10/995,561
 ; CURRENT FILING DATE: 2004-11-24
 ; NUMBER OF SEQ ID NOS: 85702
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 114
 ; LENGTH: 2443
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-995-561-114

Query Match 18.8%; Score 311.6; DB 7; Length 2443;
 Best Local Similarity 98.1%; Pred. No. 4.2e-87;
 Matches 314; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1339 GTCAGGCGCAGCAATGTTTCAGCCCGCAGTCCCGGTCGCTCAGTCAAGCAAGCCCG 1398
 DB 464 GGCAGGCTCAGCAATGTTTCAGCCCGCAGTCCCGGTCGCTCAGTCAAGCAAGCCCG 523

QY 1399 GTTGTATGACAAATGGAACCACTATCAGATAAATCAACAGTGGGAGCGGACCTACCTAG 1458
 DB 524 GTTGTATGACAAATGGAACCACTATCAGATAAATCAACAGTGGGAGCGGACCTACCTAG 583

QY 1459 GTAATGTGTGTTGTTACTTGTATGAGGAGCGGAGGTTTAACTCGGAAAGTAAAC 1518
 DB 584 GCAATGCGTGTGTTGTTACTTGTATGAGGAGCGGAGGTTTAACTCGGAGAGTAAAC 643

QY 1519 CTGAAGCTGAAGAGACTTGTCTTGAACAAGTACACTGGGAACTTACCGAGTGGGTGACA 1578
 DB 644 CTGAAGCTGAAGAGACTTGTCTTGAACAAGTACACTGGGAACTTACCGAGTGGGTGACA 703

QY 1579 CTTATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACCTGCGGCTGGGCGAG 1638
 DB 704 CTTATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACCTGCGGCTGGGCGAG 763

RESULT 11
 US-10-995-561-105
 ; Sequence 105, Application US/10995561
 ; Publication No. US20050272054A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cargill, Michele et al.
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
 ; TITLE OF INVENTION: DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001559
 ; CURRENT APPLICATION NUMBER: US/10/995,561
 ; CURRENT FILING DATE: 2004-11-24
 ; NUMBER OF SEQ ID NOS: 85702
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 105
 ; LENGTH: 2488

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2006, 16:09:44 ; Search time 6825.35 Seconds
(without alignments)
11385.988 Million cell updates/sec

Title: US-09-940-235-10
Perfect score: 1661
Sequence: 1 gcaaccgccagctagcc.....gaataagctgtaccatctaa 1661

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	312	18.8	480	2	BF993488 QVO-GN021
2	312	18.8	597	7	CN332340 170005325
3	312	18.8	657	7	CN332343 170005314
4	312	18.8	660	7	CN332334 170004554
5	312	18.8	666	7	CN332338 170006001
6	312	18.8	673	7	CN419479 170004706
7	312	18.8	676	1	AL603362 DKFZp686C
8	312	18.8	684	5	CN332332 170005326
9	312	18.8	693	8	DR004705 TC124975
10	312	18.8	706	7	CN332335 170005322
11	312	18.8	716	7	CN332342 170005339
12	312	18.8	751	1	AL706288 DKFZp686E
13	312	18.8	765	5	AX473413 DKFZp686E
14	312	18.8	911	5	AX398838 BX398838
15	312	18.8	7434	11	DQ039102 Homo sapi
16	312	18.8	7868	4	HSM806653
17	312	18.8	8411	4	HSM806992
18	311.6	18.8	1038	5	AX439175 Homo sapi
19	311	18.7	689	5	AX473425 DKFZp686L
20	311	18.7	719	5	AX473443 DKFZp686P
21	311	18.7	749	5	AX473395 DKFZp686D
22	311	18.7	759	5	AX473392 DKFZp686C

c	23	310.4	18.7	487	2	BF993485	BF993485	QVO-GN021
	24	310.4	18.7	494	3	BQ366579	BQ366579	QVO-GN021
	25	310.4	18.7	683	5	BX501294	BX501294	DKFZp779A
	26	310.4	18.7	693	1	AL706197	AL706197	DKFZp686I
	27	310.4	18.7	699	1	AL706221	AL706221	DKFZp686J
	28	310.4	18.7	748	5	BX473391	BX473391	DKFZp686C
	29	310.4	18.7	756	5	BX473407	BX473407	DKFZp686G
	30	310.4	18.7	913	5	BX380583	BX380583	BX380583
	31	310.4	18.7	1026	5	BX386270	BX386270	BX386270
	32	310.4	18.7	1050	5	BX417945	BX417945	BX417945
	33	310.4	18.7	7777	4	CR749316	CR749316	Homo sapi
	34	310.4	18.7	7885	4	CR749281	CR749281	Homo sapi
	35	310	18.7	676	5	BX510005	BX510005	DKFZp686I
	36	310	18.7	708	8	DR005445	DR005445	TC109245
	37	310	18.7	7434	11	DQ039103	DQ039103	Pan trogl
c	38	308.8	18.6	378	2	BF751799	BF751799	RC3-BN042
c	39	308.8	18.6	427	2	BF988908	BF988908	IL5-GN017
	40	308.8	18.6	8121	4	CR749317	CR749317	Homo sapi
	41	307.8	18.5	712	7	CR791105	CR791105	DKFZp469C
	42	304.4	18.3	321	1	AW068245	AW068245	cn23g07.Y
	43	303.6	18.3	925	5	BX380854	BX380854	BX380854
	44	303	18.3	670	5	BX473387	BX473387	DKFZp686B
	45	301	18.1	684	7	CN332337	CN332337	170005326

ALIGNMENTS

RESULT 1
BF993488
LOCUS BF993488 480 bp mRNA linear EST 23-JAN-2001
DEFINITION QVO-GN0216-011100-466-g03 GN0216 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF993488
VERSION BF993488.1 GI:12399811
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 480)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bata,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QVO&t=2-QVO-GN0216-011100-466-g03&t3=2000-11-01&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 24
High quality sequence stop: 479.
Location/Qualifiers
1. 480
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="GN0216"

/note="Organ: placenta normal; Vector: puc18; Site:1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 18.8%; Score 312; DB 2; Length 480;
Best Local Similarity 98.4%; Pred. No. 3e-78;
Matches 315; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1339 GTGAGCGCAGCAATGGTTTCAGCCCCCAGTCCCGGTGGCTGTGCTCAGTCAAGCAAGCCCG 1398
DB 124 GCGAGCTCAGCAATGGTTTCAGCCCCCAGTCCCGGTGGCTGTGCTCAGTCAAGCAAGCCCG 183

QY 1399 GTTGTATGCAATGAAACACTATCAGATAAATCAACAGTGGGAGCGGACCTACCTAG 1458
DB 184 GTTGTATGCAATGAAACACTATCAGATAAATCAACAGTGGGAGCGGACCTACCTAG 243

QY 1459 GTAATGTGTTGTTTGTACTTGTATGGAGGAGCCGAGGTTTAACTCGGAAGTAAAC 1518
DB 244 GCAATGCGTGTGTTTGTACTTGTATGGAGGAGCCGAGGTTTAACTCGGAGTAAAC 303

QY 1519 CTGAAGCTGAAGAGACTTGTGTTGACAAAGTACACTGGGAACACTTACCGAGTGGGTGACA 1578
DB 304 CTGAAGCTGAAGAGACTTGTGTTGACAAAGTACACTGGGAACACTTACCGAGTGGGTGACA 363

QY 1579 CTTATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACCTGCATCGGGCTGGGCGAG 1638
DB 364 CTTATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACCTGCATCGGGCTGGGCGAG 423

QY 1639 GGAGATAAGCTGTACCATC 1658
DB 424 GGAGATAAGCTGTACCATC 443

RESULT 2

CN332340

LOCUS CN332340 597 bp mRNA linear EST 16-MAY-2004

DEFINITION 1700532545117 GRN_EB Homo sapiens CDNA 5', mRNA sequence.

ACCESSION CN332340

VERSION CN332340.1 GI:47332274

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 597)
AUTHORS Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J., Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R., Lebkowski, J. and Stanton, L.W.
TITLE Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation

JOURNAL Nat. Biotechnol. 22 (6), 707-716 (2004)

PUBMED 15146197

COMMENT Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 597 Std Error: 0.00.

FEATURES

source

1..597
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue type="embryonic stem cells, embryoid bodies
derived from H1, H7 and H9 cells"

/clone lib="GRN_EB"
/note="Oligo dt primed, full-length enriched cDNA library from embryoid body outgrowths derived from hES cell lines H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free conditions."

ORIGIN

Query Match 18.8%; Score 312; DB 7; Length 597;
Best Local Similarity 98.4%; Pred. No. 3.2e-78;
Matches 315; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1339 GTGAGCGCAGCAATGGTTTCAGCCCCCAGTCCCGGTGGCTGTGCTCAGTCAAGCAAGCCCG 1398
DB 255 GCGAGCTCAGCAATGGTTTCAGCCCCCAGTCCCGGTGGCTGTGCTCAGTCAAGCAAGCCCG 314

QY 1399 GTTGTATGCAATGAAACACTATCAGATAAATCAACAGTGGGAGCGGACCTACCTAG 1458
DB 315 GTTGTATGCAATGAAACACTATCAGATAAATCAACAGTGGGAGCGGACCTACCTAG 374

QY 1459 GTAATGTGTTGTTTGTACTTGTATGGAGGAGCCGAGGTTTAACTCGGAAGTAAAC 1518
DB 375 GCAATGCGTGTGTTTGTACTTGTATGGAGGAGCCGAGGTTTAACTCGGAGTAAAC 434

QY 1519 CTGAAGCTGAAGAGACTTGTGTTGACAAAGTACACTGGGAACACTTACCGAGTGGGTGACA 1578
DB 435 CTGAAGCTGAAGAGACTTGTGTTGACAAAGTACACTGGGAACACTTACCGAGTGGGTGACA 494

QY 1579 CTTATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACCTGCATCGGGCTGGGCGAG 1638
DB 495 CTTATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACCTGCATCGGGCTGGGCGAG 554

QY 1639 GGAGATAAGCTGTACCATC 1658
DB 555 GGAGATAAGCTGTACCATC 574

RESULT 3

CN332343

LOCUS CN332343 657 bp mRNA linear EST 16-MAY-2004

DEFINITION 17005331483454 GRN_EB Homo sapiens CDNA 5', mRNA sequence.

ACCESSION CN332343

VERSION CN332343.1 GI:47332277

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 657)
AUTHORS Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J., Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R., Lebkowski, J. and Stanton, L.W.
TITLE Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation

JOURNAL Nat. Biotechnol. 22 (6), 707-716 (2004)

PUBMED 15146197

COMMENT Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 657 Std Error: 0.00.

FEATURES

source

1..657
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue type="embryonic stem cells, embryoid bodies
derived from H1, H7 and H9 cells"
/clone lib="GRN_EB"
/note="Oligo dt primed, full-length enriched cDNA library

from embryoid body outgrowths derived from hES cell lines H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free conditions."

ORIGIN

Query Match 18.8%; Score 312; DB 7; Length 657;
Best Local Similarity 98.4%; Pred. No. 3.3e-78;
Matches 315; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1339 GTCAGCGCAGCAAAATGGTTTCAGCCCGAGTCCCGGGTGGCTGTCTCAGTCAAGCAAGCCCG 1398
Db 287 GCGAGGCTCAGCAAAATGGTTTCAGCCCGAGTCCCGGGTGGCTGTCTCAGTCAAGCAAGCCCG 346
Qy 1399 GTTGTTATGACAATGAAACACTATCAGATAAATCAACAGTGGAGCGGACCTTACCTAG 1458
Db 347 GTTGTTATGACAATGAAACACTATCAGATAAATCAACAGTGGAGCGGACCTTACCTAG 406
Qy 1459 GTAATGTGTTGTTTGTACTTGTATGGAGGAGCGAGGTTTAACTGGAGAGTAAAC 1518
Db 407 GCAATGCGTTGTTTGTACTTGTATGGAGGAGCGAGGTTTAACTGGAGAGTAAAC 466
Qy 1519 CTGAAGCTGAAGAGACTTGTCTTGACAAGTACACTGGGACACTTACCGAGTGGTGACA 1578
Db 467 CTGAAGCTGAAGAGACTTGTCTTGACAAGTACACTGGGACACTTACCGAGTGGTGACA 526
Qy 1579 CTTATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTCTGCTGATCGGGGCTGGGCGAG 1638
Db 527 CTTATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTCTGCTGATCGGGGCTGGGCGAG 586
Qy 1639 GGAGAAATAGCTGTACCATC 1658
Db 587 GGAGAAATAGCTGTACCATC 606

RESULT 4

CN332334 660 bp mRNA linear EST 16-MAY-2004
LOCUS 1700045431365 GRN_EB Homo sapiens cDNA 5', mRNA sequence.
DEFINITION
ACCESSION CN332334
VERSION CN332334.1 GI:47332268
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 660)
AUTHORS Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J., Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R., Lebkowski, J. and Stanton, L.W.
TITLE Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation
JOURNAL Nat. Biotechnol. 22 (6), 707-716 (2004)
PUBMED 15146197
COMMENT Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 660 Std Error: 0.00.

FEATURES

source

1. .660
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, embryoid bodies derived from H1, H7 and H9 cells"
/clone_lib="GRN EB"
/note="oligo dt primed, full-length enriched cDNA library from embryoid body outgrowths derived from hES cell lines H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free

conditions."

ORIGIN

Query Match 18.8%; Score 312; DB 7; Length 660;
Best Local Similarity 98.4%; Pred. No. 3.3e-78;
Matches 315; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1339 GTCAGCGCAGCAAAATGGTTTCAGCCCGAGTCCCGGGTGGCTGTCTCAGTCAAGCAAGCCCG 1398
Db 325 GCGAGGCTCAGCAAAATGGTTTCAGCCCGAGTCCCGGGTGGCTGTCTCAGTCAAGCAAGCCCG 384
Qy 1399 GTTGTTATGACAATGAAACACTATCAGATAAATCAACAGTGGAGCGGACCTTACCTAG 1458
Db 385 GTTGTTATGACAATGAAACACTATCAGATAAATCAACAGTGGAGCGGACCTTACCTAG 444
Qy 1459 GTAATGTGTTGTTTGTACTTGTATGGAGGAGCGAGGTTTAACTGGAGAGTAAAC 1518
Db 445 GCAATGCGTTGTTTGTACTTGTATGGAGGAGCGAGGTTTAACTGGAGAGTAAAC 504
Qy 1519 CTGAAGCTGAAGAGACTTGTCTTGACAAGTACACTGGGAAACACTTACCGAGTGGTGACA 1578
Db 505 CTGAAGCTGAAGAGACTTGTCTTGACAAGTACACTGGGAAACACTTACCGAGTGGTGACA 564
Qy 1579 CTTATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTCTGCTGATCGGGGCTGGGCGAG 1638
Db 565 CTTATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTCTGCTGATCGGGGCTGGGCGAG 624
Qy 1639 GGAGAAATAGCTGTACCATC 1658
Db 625 GGAGAAATAGCTGTACCATC 644

RESULT 5

CN332338 666 bp mRNA linear EST 16-MAY-2004
LOCUS 17000600181351 GRN_PRENU Homo sapiens cDNA 5', mRNA sequence.
DEFINITION
ACCESSION CN332338
VERSION CN332338.1 GI:47332272
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 666)
AUTHORS Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J., Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R., Lebkowski, J. and Stanton, L.W.
TITLE Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation
JOURNAL Nat. Biotechnol. 22 (6), 707-716 (2004)
PUBMED 15146197
COMMENT Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 666 Std Error: 0.00.

FEATURES

source

1. .666
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cell, retinoic acid and mitogen-treated hES cell line H7"
/clone_lib="GRN_PRENU"
/note="oligo dt primed, full-length enriched cDNA library from hES cell line H7 (p29) maintained in feeder-free conditions. Embryoid bodies were generated in the presence of all-trans retinoic acid and mitogens."

ORIGIN

```
Query Match      18.8%; Score 312; DB 7; Length 666;
Best Local Similarity 98.4%; Pred. No. 3.3e-78;
Matches 315; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1339 GTCAGCGGAGCAAAATGGTTTCAGCCCGGAGTCCCGGTCGTGTCAGTCAAAAGCAAGCCCG 1398
Db 297 GGCAGGCTCAGCAAAATGGTTTCAGCCCGGAGTCCCGGTCGTGTCAGTCAAAAGCAAGCCCG 356
QY 1399 GTTGTTATGACAAATGGAACACTATCAGATAAATCAACAGTGGGAGCGGACCTACCTAG 1458
Db 357 GTTGTTATGACAAATGGAACACTATCAGATAAATCAACAGTGGGAGCGGACCTACCTAG 416
QY 1459 GTAATGTGTGTTGTTGTTGTTATGAGGAGCGGAGTCTTTAACTGCGAAAGTAAC 1518
Db 417 GCAATGCGTGTGTTGTTGTTGTTATGAGGAGCGGAGTCTTTAACTGCGAGTAAC 476
QY 1519 CTGAAGCTGAAGAGACTTCTTTGACAAGTACACTCGGGAACACTTACCGAGTGGGTGACA 1578
Db 477 CTGAAGCTGAAGAGACTTCTTTGACAAGTACACTCGGGAACACTTACCGAGTGGGTGACA 536
QY 1579 CTTATGAGGCTCTAAAGACTCCATGATCTGGGACTGTACCTGCGGCGTGGGGCGAG 1638
Db 537 CTTATGAGGCTCTAAAGACTCCATGATCTGGGACTGTACCTGCGGCGTGGGGCGAG 596
QY 1639 GGAGATAAGCTGTACCATC 1658
Db 597 GGAGATAAGCTGTACCATC 616

RESULT 6
CN419479      673 bp mRNA linear EST 16-MAY-2004
LOCUS      17000470666007 GRN_EB Homo sapiens cDNA 5', mRNA sequence.
DEFINITION      CN419479
ACCESSION      CN419479
VERSION      CN419479.1 GI:47407073
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE      1 (bases 1 to 673)
AUTHORS      Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,
Li, Y., Xu, C., Pang, R., Guegler, K., Rao, M.S., Mandalam, R.,
Lebkowski, J. and Stanton, L.W.
TITLE      Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation
JOURNAL      Nat. Biotechnol. 22 (6), 707-716 (2004)
PUBMED      15146197
COMMENT      Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 673 Std Error: 0.00.
FEATURES      Location/Qualifiers
source      1..673
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /tissue_type="embryonic stem cells, embryoid bodies
            derived from H1, H7 and H9 cells"
            /clone_lib="GRN EB"
            /note="oligo dt primed, full-length enriched cDNA library
            from embryoid body outgrowths derived from hES cell lines
            H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free
            conditions."

ORIGIN
Query Match      18.8%; Score 312; DB 7; Length 673;
Best Local Similarity 98.4%; Pred. No. 3.3e-78;

QY 1339 GTCAGCGGAGCAAAATGGTTTCAGCCCGGAGTCCCGGTCGTGTCAGTCAAAAGCAAGCCCG 1398
Db 193 GGCAGGCTCAGCAAAATGGTTTCAGCCCGGAGTCCCGGTCGTGTCAGTCAAAAGCAAGCCCG 252
QY 1399 GTTGTTATGACAAATGGAACACTATCAGATAAATCAACAGTGGGAGCGGACCTACCTAG 1458
Db 253 GTTGTTATGACAAATGGAACACTATCAGATAAATCAACAGTGGGAGCGGACCTACCTAG 312
QY 1459 GTAATGTGTGTTGTTGTTGTTATGAGGAGCGGAGTCTTTAACTGCGAAAGTAAC 1518
Db 313 GCAATGCGTGTGTTGTTGTTGTTATGAGGAGCGGAGTCTTTAACTGCGAGTAAC 372
QY 1519 CTGAAGCTGAAGAGACTTCTTTGACAAGTACACTCGGGAACACTTACCGAGTGGGTGACA 1578
Db 373 CTGAAGCTGAAGAGACTTCTTTGACAAGTACACTCGGGAACACTTACCGAGTGGGTGACA 432
QY 1579 CTTATGAGGCTCTAAAGACTCCATGATCTGGGACTGTACCTGCGGCGTGGGGCGAG 1638
Db 433 CTTATGAGGCTCTAAAGACTCCATGATCTGGGACTGTACCTGCGGCGTGGGGCGAG 492
QY 1639 GGAGATAAGCTGTACCATC 1658
Db 493 GGAGATAAGCTGTACCATC 512

RESULT 7
AL603362      676 bp mRNA linear EST 04-SEP-2003
LOCUS      DKFZp686C197_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
DEFINITION      DKFZp686C197_5', mRNA sequence.
ACCESSION      AL603362
VERSION      AL603362.1 GI:15166868
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE      1 (bases 1 to 676)
AUTHORS      Ansorge, W., Wirkner, U., Mewes, W., Weil, B. and Wiemann, S.
TITLE      EST (Ansorge, W., Wirkner, U., Mewes, H.W., Weil, B. and Wiemann, S.)
JOURNAL      Unpublished (1999)
COMMENT      Contact: MIPS
MIPS      Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by EMBL (European Molecular Biology Laboratories,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No sl sequence available.
This clone (DKFZp686C197) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heuberweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES      Location/Qualifiers
source      1..676
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="DKFZp686C197"
            /dev_stage="adult"
            /lab_host="DH10B"
            /clone_lib="686 (synonym: hlcc3)"
            /note="Vector: pTriplex2; Site_1: SfIIA; Site_2: SfiIB;
            cDNA-collection"

ORIGIN
Query Match      18.8%; Score 312; DB 1; Length 676;
Best Local Similarity 98.4%; Pred. No. 3.3e-78;
```


Matches 315; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1339 GTGAGCGCAGCAAAATGGTTTCAGCCCGCAGTCCCGGTGGCTGTCTCAGTCAAGCAAGCCCG 1398

Db 18 GCGAGCGCTCAGCAAAATGGTTTCAGCCCGCAGTCCCGGTGGCTGTCTCAGTCAAGCAAGCCCG 77

QY 1399 GTTGTATGACAAATGGAACAACATATCAGATAAATCAACAGTGGAGCGGACCTACCTAG 1458

Db 78 GTTGTATGACAAATGGAACAACATATCAGATAAATCAACAGTGGAGCGGACCTACCTAG 137

QY 1459 GTAATGTGTTGTTTGTACTTGTATGGAGGAGCGGAGTCTTAACTGCGAAGTAAC 1518

Db 138 GCAATGCGTGTGTTGTACTTGTATGGAGGAGCGGAGTCTTAACTGCGAAGTAAC 197

QY 1519 CTGAAGCTGAAGAGACTTGTCTTGACAAAGTACACTTGGGAACACTTACCGAGTGGTGACA 1578

Db 198 CTGAAGCTGAAGAGACTTGTCTTGACAAAGTACACTTGGGAACACTTACCGAGTGGTGACA 257

QY 1579 CTTATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACCTGCATCGGGCTGGGCGAG 1638

Db 258 CTTATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACCTGCATCGGGCTGGGCGAG 317

QY 1639 GGAGATAAGCTGTACCATC 1658

Db 318 GGAGATAAGCTGTACCATC 337

RESULT 8

CN332332

LOCUS CN332332 684 bp mRNA linear EST 16-MAY-2004

DEFINITION 17000532625776 GRN_EB Homo sapiens CDNA 5', mRNA sequence.

ACCESSION CN332332

VERSION CN332332.1 GI:47332266

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 684)

AUTHORS Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fiek,G.J., Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R., Lebkowski,J. and Stanton,L.W.

TITLE Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation

JOURNAL Nat. Biotechnol. 22 (6), 707-716 (2004)

PUBMED 15146197

COMMENT Contact: Brandenberger R
Regenerative Medicine
Genon Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@genon.com
Insert Length: 684 Std Error: 0.00.

FEATURES

source

1..684

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/tissue_type="embryonic stem cells, embryoid bodies derived from H1, H7 and H9 cells"

/clone_lib="GRN EB"

/note="Oligo dT primed, full-length enriched cDNA library from embryoid body outgrowths derived from hES cell lines H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free conditions."

ORIGIN

Query Match 18.8%; Score 312; DB 7; Length 684;

Best Local Similarity 98.4%; Pred. No. 3.3e-78;

Matches 315; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1339 GTGAGCGCAGCAAAATGGTTTCAGCCCGCAGTCCCGGTGGCTGTCTCAGTCAAGCAAGCCCG 1398

Db 312 GCGAGCGCTCAGCAAAATGGTTTCAGCCCGCAGTCCCGGTGGCTGTCTCAGTCAAGCAAGCCCG 371

QY 1399 GTTGTATGACAAATGGAACAACATATCAGATAAATCAACAGTGGAGCGGACCTACCTAG 1458

Db 372 GTTGTATGACAAATGGAACAACATATCAGATAAATCAACAGTGGAGCGGACCTACCTAG 431

QY 1459 GTAATGTGTTGTTTGTACTTGTATGGAGGAGCGGAGTCTTAACTGCGAAGTAAC 1518

Db 432 GCAATGCGTGTGTTGTACTTGTATGGAGGAGCGGAGTCTTAACTGCGAAGTAAC 491

QY 1519 CTGAAGCTGAAGAGACTTGTCTTGACAAAGTACACTTGGGAACACTTACCGAGTGGTGACA 1578

Db 492 CTGAAGCTGAAGAGACTTGTCTTGACAAAGTACACTTGGGAACACTTACCGAGTGGTGACA 551

QY 1579 CTTATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACCTGCATCGGGCTGGGCGAG 1638

Db 552 CTTATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACCTGCATCGGGCTGGGCGAG 611

QY 1639 GGAGATAAGCTGTACCATC 1658

Db 612 GGAGATAAGCTGTACCATC 631

RESULT 9

DR004705

LOCUS DR004705 693 bp mRNA linear EST 17-MAY-2005

DEFINITION TC124975 Human placenta, large insert, pCMV expression library Homo sapiens cDNA clone TC124975 5', similar to Homo sapiens fibronectin 1 (FNI), transcript variant 3, mRNA sequence.

ACCESSION DR004705

VERSION DR004705.1 GI:66264578

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 693)

AUTHORS Birkett,C., Cho,J., Gau,Y., Hamer,R., Kelly,S., Kovacs,K., Liu,L., Liu,X., Porter,J., Sachs,A., Shu,Y., Sun,Z., Wong,J., Wu,M., Zhang,X., Jay,G. and He,W.

TITLE High throughput cloning of full-length human cDNAs directly from cDNA libraries optimized for large and rare transcripts

JOURNAL Unpublished (2005)

COMMENT Contact: Kovacs, KF
High Throughput cDNA Cloning
OriGene Technologies, Inc. (www.origene.com)
6 Taft Court, Suite 100, Rockville, MD 20850, USA
Tel: 301 340 3188
Fax: 301 340 8606
Email: cDNA@origene.com

This EST submission is part of an on-going human full-length cloning project at OriGene Technologies, Inc.
Please contact OriGene for access.
OriGene Technologies, Inc.
6 Taft Ct. Suite 100
Rockville, MD 20850
Tel: (301) 340-3188
http://www.origene.com
Seq primer: pCMV6 5prime forward vector primer, OriGene Technologies Inc.
Location/Qualifiers

FEATURES

source

1..693

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="TC124975"

/tissue_type="Placenta"

/clone_lib="Human placenta, large insert, pCMV expression library"

/note="vector: pCMV6-XL4; Site_1: EcoRI; Site_2: XhoI/Sall"


```
l. 763
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKF2p686I15162"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="686 (synonym: hlcc3)"
/notes="Vector: pTriplex2; Site 1:
```

```
ORIGIN
cDNA-collection"
Query Match 18.8%; Score 312; DB 5; Length 765;
Best Local Similarity 98.4%; Pred. No. 3.4e-78;
Matches 315; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1339 GTGAGCGCAGCAATGTTTCAGCCCCAGTCCCGGTGGCTGTGTCAGTCAAGCAAGCCCG 1398
DB 359 GCGAGGCTCAGCAATGTTTCAGCCCCAGTCCCGGTGGCTGTGTCAGTCAAGCAAGCCCG 418
QY 1399 GTTGTATGCAATGGAACAACATATCAGATAAATCAACAGTGGGCGGACCTACCTAG 1458
DB 419 GTTGTATGCAATGGAACAACATATCAGATAAATCAACAGTGGGCGGACCTACCTAG 478
QY 1459 GTAATGTGTTGTTTGTACTTGTATGGAGGAAGCCGAGTGTAACTGCGAAGTAAC 1518
DB 479 GCAATGCGTTGTTTGTACTTGTATGGAGGAAGCCGAGTGTAACTGCGAAGTAAC 538
QY 1519 CTGAAGCTGAAGAGACTTCTTTGACAAGTACACTGGGAACATTTACCGAGTGGTGACA 1578
DB 539 CTGAAGCTGAAGAGACTTCTTTGACAAGTACACTGGGAACATTTACCGAGTGGTGACA 598
QY 1579 CTTATGAGCGTCTTAAGACTCCATGATCTGGGACTGTACTGTCATCGGGGCTGGGCGAG 1638
DB 599 CTTATGAGCGTCTTAAGACTCCATGATCTGGGACTGTACTGTCATCGGGGCTGGGCGAG 658
QY 1639 GGAGAAATAGCTGTACCATC 1658
DB 659 GGAGAAATAGCTGTACCATC 678

RESULT 14
BX398838
LOCUS
DEFINITION
BX398838 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1065Y122 5-PRIME, mRNA sequence.
ACCESSION
BX398838.2 GI:46878456
VERSION
BX398838
KEYWORDS
EST.
SOURCE
Homo sapiens
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE
1. (bases 1 to 911)
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
JOURNAL
On May 13, 2003 this sequence version replaced gi:30625651.
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
1413.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?e=CS0D1065B110P1&c=1413.r.
FEATURES
source
Location/Qualifiers
1..911
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1065Y122"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 18.8%; Score 312; DB 5; Length 911;
Best Local Similarity 98.4%; Pred. No. 3.6e-78;
Matches 315; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1339 GTGAGCGCAGCAATGTTTCAGCCCCAGTCCCGGTGGCTGTGTCAGTCAAGCAAGCCCG 1398
DB 323 GCGAGGCTCAGCAATGTTTCAGCCCCAGTCCCGGTGGCTGTGTCAGTCAAGCAAGCCCG 382
QY 1399 GTTGTATGCAATGGAACAACATATCAGATAAATCAACAGTGGGCGGACCTACCTAG 1458
DB 383 GTTGTATGCAATGGAACAACATATCAGATAAATCAACAGTGGGCGGACCTACCTAG 442
QY 1459 GTAATGTGTTGTTTGTACTTGTATGGAGGAAGCCGAGTGTAACTGCGAAGTAAC 1518
DB 443 GCAATGCGTTGTTTGTACTTGTATGGAGGAAGCCGAGTGTAACTGCGAAGTAAC 502
QY 1519 CTGAAGCTGAAGAGACTTCTTTGACAAGTACACTGGGAACATTTACCGAGTGGTGACA 1578
DB 503 CTGAAGCTGAAGAGACTTCTTTGACAAGTACACTGGGAACATTTACCGAGTGGTGACA 562
QY 1579 CTTATGAGCGTCTTAAGACTCCATGATCTGGGACTGTACTGTCATCGGGGCTGGGCGAG 1638
DB 563 CTTATGAGCGTCTTAAGACTCCATGATCTGGGACTGTACTGTCATCGGGGCTGGGCGAG 622
QY 1639 GGAGAAATAGCTGTACCATC 1658
DB 623 GGAGAAATAGCTGTACCATC 642

RESULT 15
DQ039102
LOCUS
DEFINITION
DQ039102 Homo sapiens FNI gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION
DQ039102
VERSION
DQ039102.1 GI:66890311
KEYWORDS
GSS.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE
1. (bases 1 to 7434)
Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,
Hubisz,M.J., Fedel-Alon,A., Tanenbaum,D.M., Civallo,D.,
White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
A Scan for Positively Selected Genes in the Genomes of Humans and
Chimpanzees
(ser) PLoS Biol. 3 (6), E170 (2005)
JOURNAL
15869325
PUBMED
2. (bases 1 to 7434)
Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,
Hubisz,M.J., Fedel-Alon,A., Tanenbaum,D.M., Civallo,D.,
White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
Direct Submission
Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment. Translation starts at the beginning of
alignment.
FEATURES
source
Location/Qualifiers
1..7434
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="2"
<1..7434
/gene="FNI"
/locus_tag="HC8611"
ORIGIN
```


GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 28, 2006, 01:47:34 ; Search time 115.543 Seconds
(without alignments)
12632.698 Million cell updates/sec

Title: US-09-940-235-10
Perfect score: 2939
Sequence: 1 gcaaccccgccagctagcc.....gaataagctgtaccatctaa 1661

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 4886326

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_n2p_model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_p/US09940235/runat_27012006_144217_27548/app_query.fasta_1.7708
-DB=A_Geneseq -QFMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09940235 @CNG_1_1_846 @runat_27012006_144217_27548 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq_21.*

1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*
9: Geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1984	67.5	413	ADL92189	Adl92189 Streptoki
2	1984	67.5	414	AAR10194	Aar10194 Streptoki
3	1984	67.5	414	AAR63120	Aar63120 Streptoki
4	1984	67.5	414	AAV24794	Aay24794 Streptoco
5	1984	67.5	414	AAW94664	Aaw94664 Streptoco
6	1984	67.5	414	AAV01556	Aay01556 Native st
7	1984	67.5	414	AAV01556	Aay01556 S. equisi
8	1984	67.5	795	AAV24797	Aay24797 Streptoki
9	1981	67.4	531	AAR11829	Aar11829 FB-FB-SK

10	1980	67.4	414	2	AAW94665	Aaw94665 Streptoco
11	1966.5	66.9	813	2	AAW21728	Aaw21728 Wild type
12	1966.5	66.9	1194	2	AAW21726	Aaw21726 Streptoki
13	1963	66.8	415	9	AE93425	Aeb93425 Streptoco
14	1958	66.6	414	2	AAW86143	Aaw86143 Streptoki
15	1958	66.6	414	3	AAW01295	Aab01295 Wild type
16	1958	66.6	440	2	AAAR12889	Aar12889 Streptoki
17	1958	66.6	483	2	AAAR12885	Aar12885 Factor Xa
18	1958	66.6	483	2	AAAR12522	Aar12522 Factor Xa
19	1958	66.6	499	2	AAAR12891	Aar12891 Streptoki
20	1958	66.6	859	2	AAAR12893	Aar12893 OmpA str
21	1954	66.5	414	2	AAAR20202	Aar20202 S.equisim
22	1951.5	66.4	415	3	AAAY50870	Aay50870 Streptoco
23	1951.5	66.4	415	3	AAAY99593	Aay99593 Streptoco
24	1947	66.2	414	2	AAW86144	Aaw86144 De-immuni
25	1947	66.2	414	3	AAAB01296	Aab01296 Altered s
26	1942	66.1	440	2	AAAR06377	Aar06377 Streptoki
27	1936.5	65.9	1194	2	AAW21724	Aaw21724 Modified
28	1936.5	65.9	1194	2	AAW21725	Aaw21725 Modified
29	1928	65.6	413	2	AAV25020	Aay25020 Streptoki
30	1928	65.6	413	5	ABB80012	Abb80012 Streptoki
31	1928	65.6	413	6	ABG74199	Abg74199 Represent
32	1924	65.5	372	2	AAAR10200	Aar10200 Streptoki
33	1916	65.2	372	2	AAAR10197	Aar10197 Streptoki
34	1907.5	64.9	371	2	AAAR10195	Aar10195 Streptoki
35	1907	64.9	374	2	AAAR10198	Aar10198 Streptoki
36	1904	64.8	414	4	AAAB74940	Aab74940 Mutant st
37	1902	64.7	384	3	AAAY84007	Aay84007 Amino aci
38	1900	64.6	800	2	AAW21723	Aaw21723 Plasmidog
39	1900	64.6	1181	2	AAW21727	Aaw21727 Streptoki
40	1898	64.6	372	3	AAAY84006	Aay84006 Amino aci
41	1887	64.2	401	3	AAAY84004	Aay84004 Amino aci
42	1887	64.2	413	3	AAAY84005	Aay84005 Amino aci
43	1879	63.9	747	2	AAAR12894	Aar12894 Met-core
44	1875	63.8	369	2	AAAR12892	Aar12892 Truncated
45	1855	63.1	440	1	AAAP50620	Aap50620 Sequence

ALIGNMENTS

RESULT 1

ADL92189

ID ADL92189 standard; protein; 413 AA.

AC ADL92189;

DT 20-MAY-2004 (first entry)

DE Streptokinase protein sequence.

KW harvesting; recombinant; host cell; N-terminal leader peptide;
KW pre-peptide; lantibiotic; post-translational modification;
KW pharmaceuticals; vaccine; immunogenic.

OS Unidentified.

PN WO2003099862-A1.

XX 04-DEC-2003.

XX 26-MAY-2003; 2003WO-NL000389.

XX 24-MAY-2002; 2002EP-00077060.

XX 07-FEB-2003; 2003US-00360101.

XX (NANO-) APPLIED NANOSYSTEMS BV.

XX Moll CN, Leenhouts CJ, Kuipers OP, Driessen AJM;

XX WPI; 2004-042770/04.

PT Harvesting a desired polypeptide produced by a recombinant host cell, for
producing pharmaceuticals, comprises selecting a recombinant nucleic acid

comprising nucleic acid fragments encoding a leader peptide and the polypeptide.

Claim 4; Page 82-83; 109pp; English.

The invention relates to a novel method for harvesting a (poly)peptide produced by a recombinant host cell. The novel method involves selecting a cell comprising a first nucleic acid encoding a leader peptide and a second nucleic acid fragment encoding the desired (poly)peptide. The first and second fragments are within the same open reading frame of the first nucleic acid and the leader peptide is functionally equivalent to an N-terminal leader peptide found with the pre-peptide of a lantibiotic. The host cells and nucleic acids are useful for producing, harvesting and post-translational modification of polypeptides. The polypeptides may be used in the production of pharmaceuticals, e.g. as antigen for vaccine or immunogenic composition. This sequence represents a polypeptide relating to the novel method of the invention.

Sequence 413 AA;

Alignment Scores:

Pred. No.: 3,93e-178 Length: 413
Score: 1984.00 Matches: 383
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 67.51% Indels: 0
Dbs: 8 Gaps: 0

US-09-940-235-10 (1-1661) x ADL92189 (1-413)

183 ATTGCTGACCTGAGTGGCTGTAGACCTGCTCATCTGTCTCAACACAGCCAAATGGTGTGTT 242
1 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
243 AGCGTGTGCTGACTGTTGAGGGGACGATCAAGACATTAGTCTTAAATTTTTGAAATC 302
21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPheGluIle 40
303 GATCTAACATCAGCAGCTGCTCATGAGGAGAAAGACAGCAAGGCTTAAGTCCAAATCA 362
41 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 60
363 AAACCATTTGCTGATGATGCGCGATGTGCATATAAAGTGTGAGAAAGCTGACTACTA 422
61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeu 80
423 AAGCTATTTCAGAACAAATTGATCGCTAACGCTCCAGTAAACGACGACTACTTTGAGGTC 482
81 LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspTyrPheGluVal 100
483 ATTGATTTTGCAGGATGCAACCATTTACTGATCGAAACGGCAAGCTCTACTTTGCTGAC 542
101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
543 AAAGATGTTGCGTACCTTGGCCGACCAACCTGCTCCAGAAATTTTGTGTAAGCGCACAT 602
121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
603 GTGCGCGTTAGACCATATAAAGAAACCAATACAAACCAAGCCGAAATCTGTGTATGTG 662
141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
663 GAATATACGTACAGTTTACTCCCTTAAACCTGTATGACGATTTTCAGACAGTCTCAAA 722
161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
723 GATACTAGCTATTCAAAACACTGCTATCGTGTGACCATCATCATCTCAAGAAATCTACTA 782
181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 200
783 GCTCAAGCACAAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTTATGACGTGAC 842
201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 220

843 TCCTCAATCGTCACTCATGACAATGACATTTCCTGACGATTTTACCAATGGATCAAGAG 902
221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
903 TTTTACTTACCGTGTAAAAATCGGAAACAGCTTTATAGGATCAATAAAAAATCTGTCGTG 962
241 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 260
963 AATGAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1022
261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280
1023 GAAAGCCGATCATCCCTTTGATCGCAGTCACCTGAAACTGTTCCACATCAAAATACGTT 1082
281 GluLysProTyrAspPropheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
1083 GATGTCGATACCAACGAATTTGTAAGAGTGACGAGCTCTTAACAGCTAGCGAACGTAAC 1142
301 AspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320
1143 TTAGACTTCAGAGATTATACGATCTCTCGTGAATAAGCTAAACTACTCTTCAACAATCTC 1202
321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
1203 GATGCTTTGCTATTATGACATATACCTTAACCTGGAAGAGTAGAGGATATACGATGAC 1262
341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
1263 ACCAAACGATATACAAACCGTTTATATGGCAAGCGACCCCAAGAGAGAGATGCTAGCTAT 1322
361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 380
1323 CATTTAGCC 1331
381 HisLeuAla 383
RESULT 2
AAR10194
ID AAR10194 standard; protein; 414 AA.
XX
AC AAR10194;
XX
DT 28-MAR-1991 (first entry)
XX
DE Streptokinase encoded by synthetic gene.
XX
KW streptokinase; thrombolytic agent; myocardial infarction.
XX
OS Synthetic.
XX
PN EP407942-A.
XX
PD 16-JAN-1991.
XX
PF 11-JUL-1989; 89JP-00179432.
XX
PR 11-JUL-1989; 89JP-00179432.
PR 27-NOV-1989; 89JP-00307957.
PR 11-APR-1990; 90JP-00096830.
XX
PA (SAKA) OTSUKA PHARM FACTOR.
XX
PI Majima E, Ogino K, Ono K, Sakata Y, Uenoyama T;
XX
DR WPI; 1991-016179/03.
DR N-PSDB; AAQ10230.
XX
PT Synthetic gene encoding streptokinase - scale, high purity prodn. of
PT streptokinase used as a thrombolytic agent.
XX
PS Claim 1; Page 51; 76pp; English.
XX

CC Streptokinase and its derivatives can be produced in large quantities
CC with high purity for use as thrombolytic agents in patients with lung
CC thrombus or myocardial infarction. See also AAR10195-R10200
XX
SQ Sequence 414 AA;
Alignment Scores:
Pred. No.: 3.93e-178 Length: 414
Score: 1984.00 Matches: 383
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 67.51% Indels: 0
DB: 2 Gaps: 0
US-09-940-235-10 (1-1661) x AAR10194 (1-414)
QY 183 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCCAATGGTGT 242
Db 1 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
QY 243 AGCGTGTGCTGCTACTGTCAGGGGACCAATCAAGACATTAGTCTTAAATTTTGAATC 302
Db 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPheGluile 40
QY 303 GATCTAACATCAAGCACTGCTCATGAGGAGAAACAGACAGCAAGCTTAACTCCAAATCA 362
Db 41 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 60
QY 363 AAACATTGCTACTGATAGTGGCGGATGTCATATAAAGCTGAGAAAGCTGACTACTA 422
Db 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 80
QY 423 AAGGCTATTCAAGAACAAATTGCTGCTTAAGCTCCACAGTAACGACGACTACTTTGAGTC 482
Db 81 LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspTrpPheGluVal 100
QY 483 ATTGATTTTCAAGCGATGCAACATTACTGTATGTCGAAACGCAAGGTCTACTTTGCTGAC 542
Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTrpPheAlaAsp 120
QY 543 AAAGATGGTTCGGTAACCTTGGCGACCAACCTGTCTCAAGAATTTTGTCTAAACGCGACAT 602
Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
QY 603 GTGCGCTTAGCATATATAAGAAACCAATCAAAACCAAGCAAGTCTGTTGATGTG 662
Db 141 ValArgValArgProTrpLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
QY 663 GAATATCTGTACAGTTTACTCCCTTAAACCTGTATGACGATTTTCAGACCAGGTCTCAAA 722
Db 161 GluTrpThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
QY 723 GATACTAAGCTATTGAAAACTAGTATCGGTGACACCATCATCTCAAGAATTTACTA 782
Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 200
QY 783 GCTCAAGCAAAAGCATTTTAAACAAACCAACCCAGCTATACGATTTATGACGTCAG 842
Db 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTrpThrIleTrpGluArgAsp 220
QY 843 TCCTCAATCGTCACATCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 902
Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
QY 903 TTTACTTACCGTGTAAATAACGGGAACCAAGCTTATAGGATCAATAAAAAATCTGGTCTG 962
Db 241 PheThrTrpArgValLysAsnArgGluGlnAlaTrpArgIleAsnLysLysSerGlyLeu 260
QY 963 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1022
Db 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTrpTrpValLeuLysLysGly 280
QY 1023 GAAAGACCGGTATGATCCCTTTGATCCAGTCACCTTGAACCTGTTCCACCAATACGTT 1082

Db 281 GluLysProTrpAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTrpVal 300
QY 1083 GATGTCGATACCAACGAATTGCTTAAAGTGCAGCAGCTCTTAAACAGCTAGCGAAGCTAAC 1142
Db 301 AspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320
QY 1143 TTAGACTTCAGAGATTTTATACGATCCTCGTATAGGCTAAACTACTCTACAACAATCTC 1202
Db 321 LeuAspPheArgAspLeuTrpAspProArgAspLysAlaLysLeuLeuTrpAsnAsnLeu 340
QY 1203 GATCCTTTTGTATATGGAATACCTTAACTGGAAGAGTAGAGGATAATCAGATGAC 1262
Db 341 AspAlaPheGlyIleMetAspTrpThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
QY 1263 ACCAACCGTATCAACCGTTTATATGCGCAAGCGACCGAGGAGAGAAAGTCTAGCTAT 1322
Db 361 ThrAsnArgIleIleThrValTrpMetGlyLysArgProGluGlyGluAsnAlaSerTrp 380
QY 1323 CATTTAGCC 1331
Db 381 HisLeuAla 383
RESULT 3
AAR63120
ID AAR63120 standard; protein; 414 AA.
XX AAR63120;
XX 16-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 16-NOV-1994 (first entry)
XX Streptokinase.
XX Streptokinase; SK; Streptococcus equisimilis; plasminogen;
KW myocardial infarction.
XX Streptococcus dysgalactiae subsp. equisimilis.
XX
FH Key Location/Qualifiers
FT Region 1..352 /note= "claim 3, see CC"
FT Region 14..414 /note= "claim 1, see CC"
FT Region 120..352 /note= "claim 3, see CC"
FT Region 244..414 /note= "claim 3, see CC"
FT Region 244..352 /note= "claim 2, see CC"
XX WO9407992-A1.
PN
XX
PD 14-APR-1994.
XX
PF 05-OCT-1993; 93WO-US009502.
XX
PR 05-OCT-1992; 92US-00956692.
PR 29-SEP-1993; 93US-00128299.
XX
PA (GEHO) GEN HOSPITAL CORP.
PA (HARD) HARVARD COLLEGE.
XX
PI Reed GL;
XX
DR WPI; 1994-135561/16.
XX
PT DNA encoding a polypeptide which binds to plasminogen and corresponds to
PT region of streptokinase - useful to detect plasminogen in a sample and to
PT treat myocardial infarction.
XX
PS Disclosure; Page 40-41; 62pp; English.


```

XX Nucleic acid comprising a sequence encoding amino acids 14-414 of
CC streptokinase, which binds to plasminogen and does not have a sequence
CC comprising amino acids 60-414 is new. The polypeptide pref. comprises
CC amino acids 244-352, 1-352, 120-352 or 244-414. (Updated on 25-MAR-2003
CC to correct PN field.) (Updated on 16-OCT-2003 to standardise OS field)
XX
SQ Sequence 414 AA;
Alignment Scores:
Pred. No.: 3,936-178 Length: 414
Score: 1984.00 Matches: 383
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 67.51% Indels: 0
DB: 2 Gaps: 0
US-09-940-235-10 (1-1661) x AAR63120 (1-414)
QY 183 ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAAATTTGGTGT 242
DB 1 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal 20
QY 243 AGCGTTGCTGCTGCTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAATC 302
DB 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 40
QY 303 GATCTAACATCAGACCTGCTCATGGAGGAAAGACAGCAAGCGTAAAGTCAAAATCA 362
DB 41 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 60
QY 363 AAACCATTTGCTGATGATGGCGGATGTACATTAACCTTGAGAAAGCTGACTACTA 422
DB 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 80
QY 423 AAGGCTATTCAAGAACTGATCGCTAACCGTCCACAGTAACGACGACTACTTTGAGTTC 482
DB 81 LysAlaIleGlnGluLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100
QY 483 ATTGATTTTGAAGCGATGCAACCAATTAAGTCAAGAAACGCAAGCTTACTTTGCTGAC 542
DB 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
QY 543 AAGATGCTTTCGTTACCTTCCGACCCCACTGCTCCAGAAATTTTGTACGCGACAT 602
DB 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
QY 603 GTGCGCGTTAGACCATATAAGAAAAAACAATACAAACCAAGCGAAATCTGTGATGTG 662
DB 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
QY 663 GAATATACGTACAGTTTACTCCCTTAAACCTGTATGACGATTTTCAAGACAGGTCTCAA 722
DB 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
QY 723 GATACTAGCTATTGAAAACACTACTATCGTGTGACCATCACATCTCAAGAAATTACTA 782
DB 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 200
QY 783 GCTCAAGCACAAAGCATTTTAAACAAAAACCAACCCAGGCTATACGATTTTAACTGAC 842
DB 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleThrGluArgAsp 220
QY 843 TCCTCAATCGTCACTCATGACAAATGACATTTTTCCGTACGATTTTACCAATGGATCAAGAG 902
DB 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
QY 903 TTTACTTACCGTGTAAAAATCGGGAACAGCTTATAGATCAATAAAAAATCTGCTGTG 962
DB 241 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 260
QY 963 AATGAAGAAATAAACCAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1022

```

```

DB 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280
QY 1023 GAAAGCCCTATCATCCCTTGTATCGCAGTCACCTGAAAGCTTCCACATCAATACGTT 1082
DB 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
QY 1083 GATGTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 1142
DB 301 AspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320
QY 1143 TTAGACTTCAGACATTTATACGATCCTCGTGATAAGCTAAACTACTCTCAACAATCTC 1202
DB 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
QY 1203 GATGCTTTGCTATTATGACATACCTTAACCTGGAAGTAGAGGATATCAGCATGAC 1262
DB 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
QY 1263 ACCAACCGTATCATACACCGTTTATATGCGCAACGACCGCAAGAGAGAATGCTAGCTAT 1322
DB 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 380
QY 1323 CATTTAGCC 1331
DB 381 HisLeuAla 383
RESULT 4
AAY24794
ID AAY24794 standard; protein; 414 AA.
XX
AC AAY24794;
XX
DT 17-OCT-2003 (revised)
DT 26-AUG-1999 (first entry)
XX
DE Streptococcus equisimilis native streptokinase.
XX
KW Streptococcus; streptokinase; fibrin-dependent plasminogen activator;
KW nsk; rsk; bacterial; blood clot; thrombotic condition;
KW myocardial infarction; venous thrombosis; pulmonary embolism;
KW cerebral thrombosis; graft thrombosis; arterial thrombosis.
XX
OS Streptococcus dysgalactiae subsp. equisimilis.
XX
PN W09931247-A1.
XX
PD 24-JUN-1999.
XX
PF 15-DEC-1998; 98WO-US026694.
XX
PR 15-DEC-1997; 97US-0069497P.
XX
PA (HARD ) HARVARD COLLEGE.
XX
PI Reed GL;
XX
DR WPI; 1999-395183/33.
DR N-PSDB; AAX80492.
XX
PT N-terminally deleted streptokinase.
XX
PS Claim 30; Page 60-61; 73pp; English.
XX
CC The present invention describes an isolated bacterial protein that
CC induces fibrin-dependent plasminogen activation in a pharmaceutical
CC composition for dissolving blood clots. Also described are: (1) a
CC composition comprising an isolated modified streptokinase, the
CC modification being removal of amino acid residues in the amino terminus;
CC (2) a method for dissolving a blood clot in a subject, comprising
CC administering to the subject a fibrin-dependent streptokinase protein; a
CC nucleic acid (1) encoding a modified bacterial streptokinase; (3) an
CC expression vector comprising (1); and (4) a host cell transformed with
CC the expression vector of (3). The pharmaceutical composition comprising a

```

CC bacterial fibrin-dependent plasminogen activator is useful for dissolving
 CC blood clots in patients with a thrombotic condition, e.g. myocardial
 CC infarction, venous thrombosis, pulmonary embolism, cerebral thrombosis,
 CC graft thrombosis and arterial thrombosis. The modified streptokinase can
 CC also be used in non-human mammals. Streptokinase activation of
 CC plasminogen is at least 10-fold, preferably 100-fold greater in the
 CC presence of fibrin than in the absence of fibrin. The modified
 CC streptokinase has at least one amino acid substitution that inactivates a
 CC substrate site for proteolytic cleavage. This reduces the rate of
 CC degradation of the streptokinase at least two-fold. The present sequence
 CC represents native streptokinase (nsk). (Updated on 17-OCT-2003 to
 CC standardise OS field)
 XX
 SQ Sequence 414 AA;

Alignment Scores:
 Pred. No.: 3,93e-178 Length: 414
 Score: 1984.00 Matches: 383
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 67.51% Indels: 0
 DB: 2 Gaps: 0

US-09-940-235-10 (1-1661) x AAY24794 (1-414)

QY 183 ATTGCTGGACCTGAGTGGCTGCTGACCGCTCCATCTGTCAACAGCCCAATTGGTGT 242
 DB 1 IleAlaGlyProGluThrPheLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
 QY 243 AGCTGCTGGTACTCTTGGGGGAGCAATCAAGACATTAGTCTTAATTTTGGAAATC 302
 DB 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPheGluIle 40
 QY 303 GATCTAACATCAGACCTGCTCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 362
 DB 41 AspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 60
 QY 363 AAACCAATTTGCTACTGATGTCGCGGATGTCACATAAATCTCGAGAAGCTGACTTACTA 422
 DB 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeu 80
 QY 423 AAGGCTATTCAAGAACAAATGATCGCTTAACGCTCCACAGTAACGACGACTTCTTGGGTC 482
 DB 81 LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspPheGluVal 100
 QY 483 ATTGATTTGCAAGCATGCAACCATATCTGATCGAAACGCGAAGTCTACTTTGCTGAC 542
 DB 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
 QY 543 AAGATGTTTCGGTACCTTGGCGCCACCTGCTCCAGAAATTTTGTACGGGACAT 602
 DB 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
 QY 603 GTGCGGTTAGACCATATAAGAAAAACCAATAACAAACGAAATCTGTTGATGTG 662
 DB 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
 QY 663 GAATATCTGTACGTTTACTCCCTTAAACCCCTGATGACGATTCAGACAGGCTCTCAAA 722
 DB 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
 QY 723 GATACTAAGCTATTGAACACTAGCTATCGGTGACACCATCATCTCAAGAAATCTACTA 782
 DB 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 200
 QY 783 GCTCAAGCAAAAGCATTTTAAACAAACCCAGGCTATACGATTTATGACGTCAC 842
 DB 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 220
 QY 843 TCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 902
 DB 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240

QY 903 TTACTTACCCTGGTTAAATAATCGGAAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 962
 DB 241 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 260
 QY 963 AATGAAGAAATAACAAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1022
 DB 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280
 QY 1023 GAAAGCGGTATGATCCCTTTGATGCGCAGTCACTTGAACCTGTTCAACCATCAATACGTT 1082
 DB 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
 QY 1083 GATGTCATACCAACGAAATTCCTAAAAAGTGGAGCTCTTAAACAGCTAGCGAAACGTAAC 1142
 DB 301 AspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320
 QY 1143 TTAGACTTCAGAGATTATACGATCCTCGTGATAGGCTAAAGCTACTCTACAACAATCTC 1202
 DB 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
 QY 1203 GATGCTTTTGGTATTATGACTATACCTTAACCTGGAAGAGTAGAGGATAATCAGCATGAC 1262
 DB 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
 QY 1263 ACCAACCGTATACATAACCGTTTATATGGCAAGCGACCCGAGGAGAGAAATGCTAGCTAT 1322
 DB 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 380
 QY 1323 CATTTAGCC 1331
 DB 381 HisLeuAla 383
 RESULT 5
 AAW94664
 ID AAW94664 standard; protein; 414 AA.
 AC AAW94664;
 DX 17-OCT-2003 (revised)
 DT 04-MAY-1999 (first entry)
 XX Streptococcus equisimilis native streptokinase.
 DE Streptococcus equisimilis H46A; streptokinase; mutant; fibrinolytic;
 KW plasma clot; hydrolysis; haemolytic Streptococcus; plasminogen; plasmin;
 KW serine protease; fibrin; blood clot; thrombolytic;
 KW vascular thromboembolytic symptom; acute myocardial infarction;
 KW fibrinolysis; resistance.
 OS Streptococcus dysgalactiae subsp. equisimilis.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 14 /note= "encoded by ACC"
 FT
 XX US5876999-A.
 XX
 PD 02-MAR-1999.
 XX
 PF 06-DEC-1995; 95US-00568393.
 XX
 PR 06-DEC-1995; 95US-00568393.
 XX
 PA (NASC-) NAT SCI COUNCIL.
 XX
 PI Wu H;
 XX
 DR WPI; 1999-189643/16.
 DR N-PSDB; AAX16632.
 XX
 PT Mutant streptokinase polypeptide - useful as plasmin-resistant
 PT thrombolytic agent.
 XX

PS Claim 1; Col 7-10; 17pp; English.

XX The present invention describes a mutant streptokinase (SK) polypeptide

CC in which at least one of the amino acids in the Pro58-Lys59-Ser60-Lys61

CC segment of the corresponding native SK is replaced by another amino acid.

CC The present sequence represents native SK. SK is a secretory protein of

CC haemolytic streptococcus able to activate human plasminogen (Hplg) to

CC plasmin (Hplm), which is a serine protease able to catalyse the

CC hydrolysis of fibrin in blood clots. The SK is useful as a thrombolytic

CC agent in the treatment of vascular thromboembolytic symptoms such as

CC acute myocardial infarction. Compared with wild-type SK, the K59E mutant

CC is more resistant to degradation by human plasmin and is more effective

CC both in acting as a fibrinolytic agent and in activating human plasminogen.

CC (Updated on 17-OCT-2003 to standardise OS field)

XX

SQ Sequence 414 AA;

Alignment Scores:

Pred. No.: 3,93e-178 Length: 414

Score: 1984.00 Matches: 383

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 67.51% Indels: 0

DB: 2 Gaps: 0

US-09-940-235-10 (1-1661) x AA094664 (1-414)

QY 183 ATTGCTGACCTGAGTGGCTGTAGACCTGCTCATCTGTCAACACAGCCAAATGGTGT 242

DB 1 IleAlaGlyProGluThrLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20

QY 243 AGCGTGTCTGCTACTGCTGAGGGGAGCAATCAGACATAGCTTAAATTTTGAATC 302

DB 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 40

QY 303 GATCTAACATCATACGCTCTCATGTGAGGAGAAACAGACAGAGCGCTTAAGTCCAAATCA 362

DB 41 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 60

QY 363 AAACATTTGCTACTGATAGTGGCGGATGTCAATATAAATTGAGAAAGCTGACTACTA 422

DB 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 80

QY 423 AAGGCTATTCAAGAACATGATGCTTAAGTCCAGTCCAGTAAACGACGACTACTTTAGGTC 482

DB 81 LysAlaIleGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100

QY 483 ATTGATTTTGCAGGATGCAACCATTTACTGATCGAAACGGCAAGTCTACTTTGCTGAC 542

DB 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120

QY 543 AAAGATGTTGCGTAACTTGGCGGACCCCACTGTCCAAAGAAATTTTGTAAAGCGGACAT 602

DB 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140

QY 603 GTGCGGTTAGACCATATAAGAAACCAATACAAACCAACGCGAAATCTGTTGATGCG 662

DB 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160

QY 663 GAATATCTGTACAGTTTACTCCCTTAAACCTGTATGACGATTTTACAGCAGTCTCAAA 722

DB 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180

QY 723 GATACCTAGCTATTGAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATACTA 782

DB 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 200

QY 783 GCTCAAGCACAAGCAATTTAAACAAACCAACCCAGGCTATACGATTTATGACGTCGAC 842

DB 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 220

QY 843 TCCTCAATGCTCAGTCATGACATGACATTTTCGTTACGATTTTACCAATGGATCAAGAG 902

DB 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240

QY 903 TTTACTTACCGTGTAAAAATCGGACAAAGCTTATAGGATCAATAAAAAATCTGGCTCG 962

DB 241 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 260

QY 963 AATGAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAAGGG 1022

DB 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysGly 280

QY 1023 GAAAGCCGATGATCCCTTTTGATCGCAGTCACCTTGAAGCTTTTCAACATCAATACGTT 1082

DB 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300

QY 1083 GATGTCGATACCAACGAAATTGCTTAAAGTAGCAGCTCTTAACAGCTAGCGAACGTAAC 1142

DB 301 AspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320

QY 1143 TTAGACTTCAGAGATTATACGATCTCTCGTGATAAGCTAAACTTACTCTCAACAATCTC 1202

DB 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340

QY 1203 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAAACTAGAGGATAATCAGATGAC 1262

DB 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360

QY 1263 ACCAACGGTATCATACCGTTTATATGGCAAGCGACCCGAGAGAGAATGCTAGCTAT 1322

DB 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 380

QY 1323 CATTTAGCC 1331

DB 381 HisLeuAla 383

RESULT 6

AA01556

ID AA01556 standard; peptide; 414 AA.

XX

AC AA01556;

XX

DT 17-OCT-2003 (revised)

DT 18-JUN-1999 (first entry)

XX

DE Native streptokinase protein sequence.

XX

XX Antigenic peptide; streptokinase; streptokinase-specific antibody;

KW thrombolytic activity; thrombolytic therapy; glomerulonephritis;

KW rheumatic fever.

OS Streptococcus dysgalactiae subsp. equisimilis.

XX

PN WO9908698-A1.

XX

PD 25-FEB-1999.

XX

PF 18-AUG-1998; 98WO-US017114.

XX

PR 18-AUG-1997; 97US-0055911P.

XX

PA (HARD) HARVARD COLLEGE.

PA (GEHO) GEN HOSPITAL CORP.

XX

PI Reed GL, Parhami-Seren B;

XX

WP1; 1999-190113/16.

XX

PT New polypeptides which bind streptokinase-specific antibodies - useful in

PT thrombolytic therapy.

XX

PS Disclosure; Page 12; 44pp; English.

XX

CC The present sequence represents a native streptokinase. The specification

CC describes a polypeptide which binds to a streptokinase-specific antibody

CC and prevents the antibody binding to native streptokinase. The
 CC specification also describes a synthetic polypeptide (PI) comprising an
 CC epitope which binds to a streptokinase-specific antibody and reduces
 CC thrombolytic activity of streptokinase. PI is used in thrombolytic
 CC therapy, and to prevent or treat glomerulonephritis and rheumatic fever.
 CC (Updated on 17-OCT-2003 to standardise OS field)

SQ Sequence 414 AA;

Alignment Scores:

Pred. No.: 3-93e-178 Length: 414
 Score: 1984.00 Matches: 383
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 67.51% Indels: 0
 DB: 2 Gaps: 0

US-09-940-235-10 (1-1661) x AAY01556 (1-414)

QY 183 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCCAATGGTGT 242
 Db 1 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
 QY 243 ACCTTGCTGCTACTGTTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTGAATC 302
 Db 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPheGluIle 40
 QY 303 GATCTAACATCAGACCTGCTCATGCGAGGAGAACACAGCAGCAAGCTTAACTCCAAATCA 362
 Db 41 AspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 60
 QY 363 AAAACATTTGCTACTGATAGTGGCGGATGTCATATAAAGCTCGAGAAGCTGACTTACTA 422
 Db 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 80
 QY 423 AAGCTATTCAAGAACATGATCGCTAAGCTAACAGTCCACAGTAAACGAGTACTTTGAGGTC 482
 Db 81 LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100
 QY 483 ATTGATTTTGCAGCGCATCCACCATTTACTGATCGAAACGGCAGGCTCTACTTTGCTGAC 542
 Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
 QY 543 AAAGATGTTTCGTACCTTCGCGACCCCAACCTGTCACAGAAATTTTGTAAAGCGACAT 602
 Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
 QY 603 GTGCGGTTAGACCATATATAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 662
 Db 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
 QY 663 GAATATCTGTACAGTTTACTCCCTTAAACCTGTAGTACGATTCAGACCGTCTCAAA 722
 Db 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
 QY 723 GATACCTAAGCTATTGAACACTAGCTATCGGTGACACCATCATCTCAAGAACTACTTA 782
 Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 200
 QY 783 GCTCAAGCACAAGCATTTTAAACAAAAACCAACCCAGGCTATACGATTTATGAACGTGAC 842
 Db 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 220
 QY 843 TCCTCAATCGTCATCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 902
 Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
 QY 903 TTTACTTACCGTGTAAATCGGAAACAGCTTATAGATCAATAAAAATCTGCTCTG 962
 Db 241 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 260
 QY 963 AATGAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1022

Db 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysGly 280
 QY 1023 GAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCCACCATCAAAATACGTT 1082
 Db 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
 QY 1083 GATGTCGATACCAACGAATTCCTTAAAGAGTGAGCAGCTCTTTAACACAGCTAGCGAACGTAAC 1142
 Db 301 AspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320
 QY 1143 TTACACTTCAGAGATTTATACGATCCCTGCTGATAGGCTAAACTCTTACACAAATCTC 1202
 Db 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
 QY 1203 GATGCTTTTGGTATTATGGACTATACCTTAACCTGCGAAAGTAGAGGATAATCAGATGAC 1262
 Db 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
 QY 1263 ACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAGGAGAGAGAAATGCTAGCTAT 1322
 Db 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 380
 QY 1323 CATTTAGCC 1331
 Db 381 HisLeuAla 383
 RESULT 7
 AAY90282
 ID AAY90282 standard; protein; 414 AA.
 XX AC AAY90282;
 XX XX
 DT 12-SEP-2003 (revised)
 DT 13-OCT-2000 (first entry)
 XX DE
 XX XX
 KW Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
 KW plasminogen; human; fibrinectin; thrombolytic therapy;
 KW cardiovascular disorder; fibrinectin.
 XX OS
 XX Streptococcus dysgalactiae subsp. equisimilis.
 XX EP1024192-A2.
 XX XX
 PD 02-AUG-2000.
 XX XX
 PF 23-DEC-1999; 99EP-00310541.
 XX XX
 PR 24-DEC-1998; 98IN-DE003825.
 XX XX
 PA (COUL) CSIR COUNCIL SCI IND RES.
 XX XX
 PI Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
 PI Yadav M;
 XX XX
 DR WPI; 2000-516032/47.
 DR N-PSDB; AAA37633.
 XX XX
 PT Hybrid streptokinase-fibrin binding domain polypeptides useful for
 PT thrombolytic therapy comprises a streptokinase fused with fibrin binding
 PT domains of human fibrinectin.
 XX XX
 PS Example 3; Fig 3; 58pp; English.
 XX XX
 CC This sequence represents the human Streptococcus equisimilis
 CC streptokinase protein sequence. The invention relates to a hybrid
 CC plasminogen activator (PA) comprises a polypeptide fusion between
 CC streptokinase (SK), which are capable of plasminogen (PG) activation, and
 CC fibrin binding regions of human fibrinectin, which are from fibrin
 CC binding domains (PBD) 4 and 5 or 1 and 2. The hybrid PA possesses the
 CC ability to bind with fibrin independently and also characteristically
 CC retains a PG activation ability which becomes evident only after a

CC pronounced duration, or lag, after exposure of the PA to a suitable
 CC animal or human PG. The hybrid streptokinase-fibrin binding domain
 CC polypeptides are useful in thrombolytic therapy for various kinds of
 CC cardiovascular disorders. The hybrids have enhanced fibrin selectivity as
 CC well as kinetics of plasminogen activation that are distinct from that of
 CC natural streptokinase in being characterised by a temporary delay, or lag
 CC of several minutes in the natural rate of the catalytic conversion of
 CC plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins
 CC can bind tightly with fibrin in blood clots soon after introduction into
 CC the vascular system without significantly activating the circulating
 CC blood plasminogen to plasmin, thus aiding in the localisation of the
 CC plasminogen activation process to the site of pathological thrombus. This
 CC overcomes systemic plasminogen activation encountered during clinical use
 CC of streptokinase. (Updated on 12-SEP-2003 to standardise OS field)
 XX
 SQ Sequence 414 AA;

Alignment Scores:
 Pred. No.: 3,93e-178 Length: 414
 Score: 1984.00 Matches: 383
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 67.51% Indels: 0
 DB: 3 Gaps: 0

US-09-940-235-10 (1-1661) x AAY90282 (1-414)

QY 183 ATTGCTGACCTGAGTGGTGTGATGACCGTCTCATCTGTCACACAGCCAAATGGTGT 242
 Db 1 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
 QY 243 AGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 302
 Db 21 SerValAlaGlyThrValGluGlyThrAenGlnAspPileSerLeuLysPheGluIle 40
 QY 303 GATCTAACATCAGCACTGCTCATGGAGGAAAGACAGACAGCGTCTTAAGTCCAAATCA 362
 Db 41 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 60
 QY 363 AAACATTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 422
 Db 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLysLeuGluLysAlaAspLeuLeu 80
 QY 423 AAGGCTATTCAAGAACATTGATCGCTTAACGTCACAGTAAACGACGACTACTTGGCTC 482
 Db 81 LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100
 QY 483 ATTGATTTTGCAGCGATGCAACCACTTACTGATCGAAACGCGCAAGGCTCTACTTGGCTGAC 542
 Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
 QY 543 AAAGATGTTGCGTAACTTGGTCCGACCCCACTGTCACAGAAATTTTGTGTAAGCGGACAT 602
 Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
 QY 603 GTGCGGTTAGACCATATAAGAAACCAACCAATACAAACCAACCGCAATCTGTTGATGTCG 662
 Db 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
 QY 663 GAATATCTGTACAGTTTACTCCCTTAAACCTGTGATGACGATTTTACAGCAGGTCTCAAA 722
 Db 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
 QY 723 GATACCTAGCTATTGAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATTA 782
 Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 200
 QY 783 GCTCAAGCAACAGCATTTTAAACAAACCAACCCCGCTATACGATTTTATGAACTGAC 842
 Db 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 220
 QY 843 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTCACGATTTTACCAATGGATCAAGAG 902

Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
 QY 903 TTTTACTTACCGTCTTAAATAATCGGAACAAAGCTTATAGGATCAATAAAAAATCTGGTCTG 962
 Db 241 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 260
 QY 963 AATGAAGAAATAAACCACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGG 1022
 Db 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280
 QY 1023 GAAAGCGGTATGATCCCTTTGATCGAGTCACTTGAAACTGTTCCACATCAATACGTT 1082
 Db 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
 QY 1083 GATGTCGATCAACCAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 1142
 Db 301 AspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320
 QY 1143 TTAGACTTCAGAGATTTTATACGATCTCTGCTGATGAAAGCTAAACTTACTCTACAAATCTC 1202
 Db 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
 QY 1203 GATGCTTTGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1262
 Db 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
 QY 1263 ACCAAGCGTATCATACCGTCTTATATGCGCAAGCGCCGAGGAGAGATGCTAGCTAT 1322
 Db 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 380
 QY 1323 CATTTAGGCC 1331
 Db 381 HisLeuAla 383

RESULT 8

AAY24797
 ID AAY24797 standard; protein; 795 AA.
 XX
 AC AAY24797;
 XX
 DT 26-AUG-1999 (first entry)
 DE Streptokinase and maltose binding protein fusion protein.
 XX
 KW Streptococcus; streptokinase; fibrin-dependent plasminogen activator;
 KW mSK; rSK; bacterial; blood clot; thrombotic condition;
 KW myocardial infarction; venous thrombosis; pulmonary embolism;
 KW cerebral thrombosis; graft thrombosis; arterial thrombosis.
 XX
 OS Streptococcus dysgalactiae subsp. equisimilis.
 OS Synthetic.
 XX
 PN W09931247-A1.
 PD 24-JUN-1999.
 XX
 PF 15-DEC-1998; 98WO-US026694.
 XX
 PR 15-DEC-1997; 97US-0069497P.
 XX
 XX (HARD) HARVARD COLLEGE.
 XX
 PA Reed GL;
 PI
 XX
 DR WPI; 1999-395183/33.
 DR N-PSDB; AAX80497.
 XX
 PT N-terminally deleted streptokinase.
 XX
 PS Example; Page 48-51; 73pp; English.
 XX
 CC The present invention describes an isolated bacterial protein that
 CC induces fibrin-dependent plasminogen activation in a pharmaceutical

CC composition for dissolving blood clots. Also described are: (1) a
 CC composition comprising an isolated modified streptokinase, the
 CC modification being removal of amino acid residues in the amino terminus;
 CC (2) a method for dissolving a blood clot in a subject, comprising
 CC administering to the subject a fibrin-dependent streptokinase protein; a
 CC nucleic acid (1) encoding a modified bacterial streptokinase; (3) an
 CC expression vector comprising (1); and (4) a host cell transformed with
 CC the expression vector of (3). The pharmaceutical composition comprising a
 CC bacterial fibrin-dependent plasminogen activator is useful for dissolving
 CC blood clots in patients with a thrombotic condition, e.g. myocardial
 CC infarction, venous thrombosis, pulmonary embolism, cerebral thrombosis,
 CC graft thrombosis and arterial thrombosis. The modified streptokinase can
 CC also be used in non-human mammals. Streptokinase activation of
 CC plasminogen is at least 10-fold, preferably 100-fold greater in the
 CC presence of fibrin than in the absence of fibrin. The modified
 CC streptokinase has at least one amino acid substitution that inactivates a
 CC substrate site for proteolytic cleavage. This reduces the rate of
 CC degradation of the streptokinase at least two-fold. The present sequence
 CC represents a streptokinase and maltose binding protein fusion protein
 CC from an example of the present invention
 XX
 SQ Sequence 795 AA;

Alignment Scores:

Pred. No.: 4,98e-178 Length: 795
 Score: 1984.00 Matches: 383
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 67.51% Indels: 0
 DB: 2 Gaps: 0

US-09-940-235-10 (1-1661) x AAY24797 (1-795)

QY 183 ATGCTGGACCTGAGTGGCTGTAGACCGTCCATCTCTCAACACAGCCCAATTGGTGT 242
 DB 382 IIEAlaGlyProGluThrLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal 401
 QY 243 ACGGTGTCTGCTACTGTGAGGGGGAATCAACACATAGTCTTAATTTTGAATC 302
 DB 402 SerValAlaGlyThrValGluGlyThrAsnGlnAspSerLeuLysPheGluLe 421
 QY 303 GATCTAACATCAGACCTGCTCATGGAGGAAAGACAGCAAGCTTAAAGTCCAAATCA 362
 DB 422 AspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 441
 QY 363 AAACCATTTGCTACTGATAGTGGCGGATGTCATAACTTGAGAAAGCTGACTTACTA 422
 DB 442 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 461
 QY 423 AAGGCTATTTCAGAACAAATGATCGCTAAAGTCCACAGTAACGACGACTACTTGGGTC 482
 DB 462 LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 481
 QY 483 ATTGATTTTTCAGCGCATGCAACCATTTACTGATCGAAACGCAAGGCTACTTTGCTGAC 542
 DB 482 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 501
 QY 543 AAAGATGGTTCGGTAACCTTTCGCGACCCCAACCTGTCCAAGAATTTTGTGAAGCGACAT 602
 DB 502 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 521
 QY 603 GTGCGGTTAGACCATATAAGAAAAAACCAATACAAACCAAGCGAAATCTGTGTATGTG 662
 DB 522 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 541
 QY 663 GAATATACGTGACGTTTACTCCCTTAAACCCGTAGTACGATGCGATTTTCAGACGCTCTCAA 722
 DB 542 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 561
 QY 723 GATCTAAGCTATTGAAAAACCTAGCTTATCGGTGACACCATCATCATCTCAAGAAATTACTA 782
 DB 562 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 581

QY 783 GCTCAAGCACAAAGCATTTTAAACAAAAACCAACCCAGGCTATACGATTTATGAACGTGAC 842
 DB 582 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgasp 601
 QY 843 TCCTCAATCGTCACCTCATGACAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 902
 DB 602 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 621
 QY 903 TTTTACTTACCTGTTTAAAAATCGGCAACAGCTTATAGTATCAATAAAAAATCTGCTGTG 962
 DB 622 PheThrTyrArgValLysAsnArgGlnAlaTyrArgIleAsnLysLysSerGlyLeu 641
 QY 963 AATGAAGAAATAAAACAAACACCTGATCTCTGAGAAATATTAGCTCTTAAAAAAGGG 1022
 DB 642 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 661
 QY 1023 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCCACCATCAAAATACGTT 1082
 DB 662 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 681
 QY 1083 GATGTCGATACCAAGCAATTCCTTAAAGATGAGCAGCTCTTAAACAGCTAGCGAAGCTAAC 1142
 DB 682 AspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 701
 QY 1143 TTAGACTTCAGAGATTTATACGATCCTCGTATAGGCTTAAACTACTCTACAAATCTC 1202
 DB 702 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 721
 QY 1203 GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAAGTAGAGGATAATCAGCATGAC 1262
 DB 722 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 741
 QY 1263 ACCAACCGTATCATAAACCGTTTATATGGCAAGCAGCCGCAAGAGAGAGATGCTAGCTAT 1322
 DB 742 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 761
 QY 1323 CATTTAGCC 1331
 DB 762 HisLeuAla 764
 RESULT 9
 AAR11829
 ID AAR11829 standard; protein; 531 AA.
 XX
 AC AAR11829;
 DT 08-JUL-1991 (first entry)
 XX
 DE FB-FB-SK fusion conjugate.
 KW Fibrin-binding protein; fibrinolysis; intravascular thrombi; fibrinogen;
 KW streptokinase; fusion protein.
 XX
 OS Staphylococcus aureus.
 XX
 FH Key Location/Qualifiers
 FT Peptide 2..61
 FT /label= FB monomer
 FT Peptide 62..119
 FT /label= FB monomer
 FT Peptide 120..531
 FT /label= streptokinase
 XX
 FN US5011686-A.
 XX
 PD 30-APR-1991.
 XX
 PF 15-NOV-1989; 89US-00437769.
 XX
 PR 21-SEP-1987; 87US-00099242.
 XX
 PA (CREA-) CREATIVE BIOMOLEC.
 XX

PI Pang RHL;
 XX WPI; 1991-140198/19.
 DR N-PSDB; AAQ11651.
 XX
 PT Imparting injectable fibrinolytic agent - with affinity for intravascular
 PT thrombus, by linking agent to fibrin binding domain.
 XX
 PS Disclosure; Fig 5; 18pp; English.
 XX
 CC The conjugate comprises an PB-FB dimer linked to streptokinase The PB
 CC fragment has selective affinity for fibrin, low affinity for fibrinogen,
 CC and minimal immunogenicity, imparting thrombus- targeting capability.
 CC See also AAR11821 and AAR11828
 XX
 XX
 SQ Sequence 531 AA;
 Alignment Scores:
 Pred. No.: 8.26e-178 Length: 531
 Score: 1981.00 Matches: 382
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 99.74% Mismatches: 0
 Query Match: 67.40% Indels: 0
 DB: 2 Gaps: 0
 US-09-940-235-10 (1-1661) x AAR11829 (1-531)
 QY 183 ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCCCAATGGTGT 242
 DB 118 IleaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnThrGlnLeuVal 137
 QY 243 AGCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 302
 DB 138 ServAlaGlyThrValGluGlyThrAsnGlnAspPheLeuLeuPheGluile 157
 QY 303 GATCTAACATCAGCAGCTCTCATGAGGAGAAACAGACAGCAAGCTTAAGTCAAAATCA 362
 DB 158 AspLeuThrSerArgProAlaHisGlyGlyLeuThrGluGlnGlyLeuSerProLysSer 177
 QY 363 AAACATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 422
 DB 178 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 197
 QY 423 AAGCTATTCAAGCAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 482
 DB 198 LysAlaIleGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 217
 QY 483 ATTGATTTTGCAAGCGATGCAACCATTTACTGATCGAAACCGCAAGGCTACTTTGCTGAC 542
 DB 218 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 237
 QY 543 AAGATAGTTCGGTAACTTCGCGACCCCACTGCTCAAGAATTTTTCGCTTAACGCGACAT 602
 DB 238 LysAspGlySerValThrLeuProThrGlnProValGlnGlnPheLeuLeuSerGlyHis 257
 QY 603 GTGCGGTTAGACCATATAAGAAACCAATACAAACCAACCAACCAACCAACCAACCAAC 662
 DB 258 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 277
 QY 663 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTCAGACGAGTCTCAAA 722
 DB 278 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 297
 QY 723 GATACCTAGCTATTGAACACCTAGCTATCGGTGACACCATCATCATCTCAAGAAATCTA 782
 DB 298 AspThrLysLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 317
 QY 783 GCTCAAGCAAGCAATTTTAAACAAACCAACCAACCAACCAACCAACCAACCAACCAAC 842
 DB 318 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 337
 QY 843 TCCTCAATCGTCACTCATGCAATGACATTTTCGTTACGATTTTACCAATGGAATCAAGAG 902

DB 338 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 357
 QY 903 TTTACTTTACCGTGTAAAAATCGGAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 962
 DB 358 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 377
 QY 963 AATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAGGG 1022
 DB 378 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 397
 QY 1023 GAAAGCCGATGATCCCTTTGATCGAGTCACCTTGAACACTGTTCAACATCAATACGTT 1082
 DB 398 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 417
 QY 1083 GATCTCGATACCAACGAATTTGCTAAAGAGTGAAGGAGCTCTTAACAGCTAGCGAACGTAAC 1142
 DB 418 AspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 437
 QY 1143 TTAGACTTCAGAGATTTTATACGATCCTCGTGATGAAGGCTAAAGCTTACTCTACAACTCTC 1202
 DB 438 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 457
 QY 1203 GATGCTTTTGGTATTATGACTATACCTTAACCTGGAAGAGTAGAGGATAATCAGATGAC 1262
 DB 458 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 477
 QY 1263 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCGAGGAGAGATGCTAGCTAT 1322
 DB 478 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 497
 QY 1323 CATTTAGCC 1331
 DB 498 HisLeuAla 500
 RESULT 10
 AAW94665
 ID AAW94665 standard; protein; 414 AA.
 XX
 AC AAW94665;
 XX
 DT 04-MAY-1999 (first entry)
 XX
 DE Streptococcus equisimilis mutant streptokinase K59E.
 XX
 KW Streptococcus equisimilis H46A; streptokinase; mutant; fibrinolytic;
 KW plasma clot; hydrolysis; haemolytic streptococcus; plasminogen; plasmin;
 KW serine protease; fibrin; blood clot; thrombolytic;
 KW vascular thromboembolytic symptom; acute myocardial infarction;
 KW fibrinolysis; resistance.
 XX
 OS Streptococcus dysgalactiae subsp. equisimilis.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 14 /note= "encoded by ACC"
 XX
 PN US5876999-A.
 XX
 PD 02-MAR-1999.
 XX
 PF 06-DEC-1995; 95US-00568393.
 XX
 PR 06-DEC-1995; 95US-00568393.
 XX
 PA (NASC-) NAT SCI COUNCIL.
 XX
 PI Wu H;
 XX
 DR WPI; 1999-189643/16.
 DR N-PSDB; AAX16633.
 XX
 PT Mutant streptokinase polypeptide - useful as plasmin-resistant

thrombolytic agent.

Claim 4; Col 11-14; 17pp; English.

The present invention describes a mutant streptokinase (SK) polypeptide in which at least one of the amino acids in the Pro58-Lys59-Ser60-Lys61 segment of the corresponding native SK is replaced by another amino acid. The present sequence is mutant SK K59E. SK is a secretory protein of haemolytic Streptococcus able to activate human plasminogen (HPLg) to plasmin (HPLm), which is a serine protease able to catalyse the hydrolysis of fibrin in blood clots. The SK is useful as a thrombolytic agent in the treatment of vascular thromboembolytic symptoms such as acute myocardial infarction. Compared with wild-type SK, the K59E mutant is more resistant to degradation by human plasmin and is more effective both in acting as a fibrinolytic agent and in activating human plasminogen

Sequence 414 AA;

Alignment Scores:

Pred. No.:	9.38e-178	Length:	414
Score:	1980.00	Matches:	382
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.74%	Mismatches:	0
Query Match:	67.37%	Indels:	0
DB:	2	Gaps:	0

US-09-940-235-10 (1-1661) x AAW94665 (1-414)

QY	183	ATTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCACACAGCCAAATGGTGT	242
DB	1	IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal	20
QY	243	AGCGTTGCTGTACTGCTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAATC	302
DB	21	SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle	40
QY	303	GATCTAACATCACGACCTGCTCATGGAGGAAGACAGACGAGCTTAAGTCCAAATCA	362
DB	41	AspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProGluSer	60
QY	363	AAACCATTTGCTACTGATAGTGGCGGATGTCATAAACTTGAGAAAGCTGACTACTA	422
DB	61	LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu	80
QY	423	AAGGCTATTCAAGACAAATGTGCTTAACGTCACAGTCCACAGTAAACGACTACTTTGAGTC	482
DB	81	LysAlaIleGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal	100
QY	483	ATTGATTTTGCAGGATGCAACCATTTACTGATCGAAACGGCAAGTCTACTTTGCTGAC	542
DB	101	IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp	120
QY	543	AAAGATGTTGCGTAACTTGGCGGACCCCAACCTGTCACAGAAATTTTGTAAAGCGGACAT	602
DB	121	LysAspGlySerValThrLeuProThrGlnProValGlnGlnPheLeuLeuSerGlyHis	140
QY	603	GTGCGGTTAGACCATATAAAGAAAACCAATACAAAACCAAGCGAAATCTGTGTATG	662
DB	141	ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal	160
QY	663	GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCGGTCTCAA	722
DB	161	GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys	180
QY	723	GATACCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATTA	782
DB	181	AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu	200
QY	783	GCTCAAGCACAAAGCATTTAAACAAAACCAACCCAGGCTATACGATTTATGACGTGAC	842
DB	201	AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp	220
QY	843	TCCTCAATCGTCACTCATGACAAATGCATTTTCCGTACGATTTTACCATGGATCAAGAG	902

DB	221	SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu	240
QY	903	TTTACTTACCGTGTAAATAATCGGGAACAAGCTTATAGATCAATAAAATCTGGTCTG	962
DB	241	PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu	260
QY	963	AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG	1022
DB	261	AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly	280
QY	1023	GAAGAAGCGTATGATCCCTTTGATCGGAGTCACCTTGAAACTGTTCCACCAATACGTT	1082
DB	281	GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal	300
QY	1083	GATGCTCATACCAACGAATTGCTAAAAGTCAGCAGCTCTTAAACAGCTAGCGAAGCTAAC	1142
DB	301	AspValAspThrAsnGluLeuLeuLysSerGluGlnLeuThrAlaSerGluArgAsn	320
QY	1143	TTAGACTTCAGAGATTTATACGATCCCTGATAGAGCTAAACTACTCTACAACTCTC	1202
DB	321	LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuTyrAsnAsnLeu	340
QY	1203	GATGCTTTTGGTATTATGGAATATACCTTAACTCGAAAAGTAGAGGATAATCACGATGAC	1262
DB	341	AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp	360
QY	1263	ACCAACCGTATCATACCGTTTATATGGGCAAGGACCCGAGGAGAGAATGCTAGCTAT	1322
DB	361	ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr	380
QY	1323	CATTTAGGCC 1331	
DB	381	HisLeuAla 383	

RESULT 11

AAW21728
ID AAW21728 standard; protein; 813 AA.

AC AAW21728;

XX 17-OCT-2003 (revised)

DT 01-OCT-1997 (first entry)

XX Wild type plasminogen-binding fragment of Streptokinase.

XX Plasminogen-binding fragment; streptokinase; degradation; MBP;

XX thrombolytic agent; blood clot; bolus; maltose-binding protein.

OS Streptococcus dysgalactiae subsp. equisimilis.

XX WO9641883-A1.

XX 27-DEC-1996.

XX 07-JUN-1996; 96WO-US009640.

XX 09-JUN-1995; 95US-00488940.

XX (HARD) HARVARD COLLEGE.

XX Reed GL;

XX WPI; 1997-065469/06.

XX Modified forms of streptokinase resistant to enzymatic cleavage - useful as thrombolytic agents in treating thrombosis and in medical equipment.

XX Example 1; Page 12-13; 65pp; English.

XX This sequence represents the wild type plasminogen-binding fragment of streptokinase. This fragment was used in the design of a modified streptokinase has an in vitro degradation rate at least 2 times slower

CC than that of native streptokinase. Compounds containing modified
 CC streptokinases are specifically used as thrombolytic agents for
 CC dissolving blood clots in vivo in a mammal, preferably at a dose of 20000
 CC U/kg, opt. as a bolus rather than by continuous infusion. (Updated on 17-
 CC OCT-2003 to standardise OS field)
 XX

SQ Sequence 813 AA;

Alignment Scores:

Pred. No.: 2,266-176 Length: 813
 Score: 1966.50 Matches: 382
 Percent Similarity: 99.74% Conservative: 0
 Best Local Similarity: 99.74% Mismatches: 0
 Query Match: 66.91% Indels: 1
 DB: 2 Gaps: 1

US-09-940-235-10 (1-1661) x AAW21728 (1-813)

Qy 183 ATTGCTGGACCTGAGTGGCTGTAGACCGTCTCAACACAGCCCAATTTGGTTGT 242
 Db 1 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal 20
 Qy 243 AGCGTTGCTGTACTGTTGAGGGGACGAATCAAGACATAGTCTTAATTTTGAATC 302
 Db 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPheGluIle 40
 Qy 303 GATCTAAACATCAGACCTCTCATGAGGAGAAAGACAGCAAGCTTAAGTCCAAATCA 362
 Db 41 AspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 60
 Qy 363 AAACCATTTGCTACTGATAGTGGCGGATGTCATATAACTTCAGAAAGCTGACTTACTA 422
 Db 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 80
 Qy 423 AAGGTATTCAAGACAAATTGATCGCTAAGTAAAGTCCACAGTAACGACGACTACTTTGAGTC 482
 Db 81 LysAlaIleGlnGluLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100
 Qy 483 ATTGATTTTGCAGGATCAACCATTAAGTATGATCGAAACGGCAAGTCTACTTTGCTGAC 542
 Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
 Qy 543 AAAGATGTTTCGGTAACCTTCGGACCCCAACCTGTCACAGAAATTTTGTCTAAGCGACAT 602
 Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
 Qy 603 GTGCGCGTTAGACCATATAAAGAAACCAATACAAACCAACGAAATCTGTGTATGTG 662
 Db 141 ValArgValArg--TyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 159
 Qy 663 GAATATCTGTACAGTTTACTTCCTTAAACCTGTATGACGATTCAGACAGGTCTCAAA 722
 Db 160 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 179
 Qy 723 GATCTAAGCTATTGAACACTAGTATCGTGGTGACACCATCATCATCTCAAGAATTACTA 782
 Db 180 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 199
 Qy 783 GCTCAAGCACAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTATGAACGTGAC 842
 Db 200 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 219
 Qy 843 TCCTCAATGTCATCATGACAAATGACATTTTCGTTACGATTTTACCAATGGATCAAGAG 902
 Db 220 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 239
 Qy 903 TTACTACTCCGTTTAAATCGGACCAAGCTTATAGGATCAATAAAATCTGCTCTG 962
 Db 240 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 259
 Qy 963 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1022
 Db 260 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 279

Qy 1023 GAAAAGCGGTATGATCCCTTTTGTGCGAGTCACTTGAACCTGTTCCACCATCAAAATACGTT 1082
 Db 280 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 299
 Qy 1083 GATGTCGATACCAACGAATTGCTTAAAGTGAGCAGCTCTTAACAGCTAGCGAACGTAAC 1142
 Db 300 AspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 319
 Qy 1143 TTAGACTTCAGAGATTTTATACGATCCCTCGTCATAGGCTAACTACTCTACACAAATCTC 1202
 Db 320 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 339
 Qy 1203 GATGCTTTTGGTATTATGGACTATACCTTAACCTCGAAAGTAGAGGATAATCAGCATCAC 1262
 Db 340 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 359
 Qy 1263 ACCAACGCTATCATACCGTTTATATGGCAAGCGACCCGAGGAGAGAATCTAGTATAT 1322
 Db 360 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 379
 Qy 1323 CATTTAGCC 1331
 Db 380 HisLeuAla 382

RESULT 12

AAW21726

ID AAW21726 standard; protein; 1194 AA.

AC AAW21726;

DT 17-OCT-2003 (revised)

DT 01-OCT-1997 (first entry)

XX Streptokinase/maltose binding protein fusion protein, rSK.

XX Plasminogen-binding fragment; streptokinase; degradation; MBP;
 KW thrombolytic agent; blood clot; bolus; maltose-binding protein.

OS Streptococcus dysgalactiae subsp. equisimilis.

XX Key Location/Qualifiers

FT Protein 1..381

FT /label= Maltose binding protein

FT /note= "acts as blocking group"

FT 382..1194

FT /label= Streptokinase

XX WO9641883-A1.

XX 27-DEC-1996.

XX 07-JUN-1996; 96WO-US009640.

XX 09-JUN-1995; 95US-00488940.

XX (HARD) HARVARD COLLEGE.

XX Reed GL;

XX WPI; 1997-065469/06.

XX Modified forms of streptokinase resistant to enzymatic cleavage - useful
 as thrombolytic agents in treating thrombosis and in medical equipment.

PS Example 1; Page 11-12; 65pp; English.

XX This sequence represents a fusion protein between maltose-binding protein
 CC and the plasminogen-binding fragment of streptokinase. This fusion
 CC protein was used in the design of modified streptokinase has an in vitro
 CC degradation rate at least 2 times slower than that of native
 CC streptokinase. Compounds containing modified streptokinases are
 CC specifically used as thrombolytic agents for dissolving blood clots in

CC vivo in a mammal, preferably at a dose of 20000 U/kg, optionally as a
CC bolus rather than by continuous infusion. (Updated on 17-OCT-2003 to
CC standardise OS field)
XX
SQ Sequence 1194 AA;

Alignment Scores:
Pred. No.: 2,6e-176 Length: 1194
Score: 1966.50 Matches: 382
Percent Similarity: 99.74% Conservative: 0
Best Local Similarity: 99.74% Mismatches: 0
Query Match: 66.91% Indels: 1
DB: 2 Gaps: 1

US-09-940-235-10 (1-1661) x AAW21726 (1-1194)

QY 183 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCCAATTGGTTGTT 242
Db 382 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal 401
QY 243 AGCGTTGCTGCTACTGTTGAGGGGACCAATCAAGACATTAGTCTTAAATTTTTTGAATC 302
Db 402 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 421
QY 303 GATCTAACATCAGCACTGCTCTCATGGAGGAAACACAGACAGCAAGCTTAACTCAAAATCA 362
Db 422 AspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 441
QY 363 AAAACATTGCTACTGATAGTGGCGGATGTCATAAACTTGAGAAAGCTGACTTACTA 422
Db 442 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 461
QY 423 AAGGCTATTCAAGACAATTGATCGCTTAACGTCACAGTAACACGACTACTTTGGTTC 482
Db 462 LysAlaIleGlnGluLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 481
QY 483 ATTGATTTTGAAGCATTGCAACCACTTACTGATCGAAACGGCAAGGTCTACTTTGCTGAC 542
Db 482 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 501
QY 543 AAAGATGGTTCGGTAACCTTGGCGACCCCAACCTGTCCAAGAATTTTGTCTAAACGGACAT 602
Db 502 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 521
QY 603 GTGCGGCTTAGCATTATAAGAAACCAATACAAAACCAACGAAATCTGTTGATGTG 662
Db 522 ValArgValArg---TyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 540
QY 663 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCGTCTCAA 722
Db 541 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 560
QY 723 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCATCTCAAGAATTACTA 782
Db 561 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 580
QY 783 GCTCAAGCAAAAGCAATTTTAAACAAAACCCAGCTATACGATTTATGAAACGTGAC 842
Db 581 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 600
QY 843 TCCTCAATCGTCATCATGACATGACATTTTCGTTACGATTTTACCAATGGATCAAGAG 902
Db 601 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 620
QY 903 TTTACTTACCGTGTAAAAATCGGAAACCAAGCTTATAGGATCAATAAAAAATCTGCTCTG 962
Db 621 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 640
QY 963 AATGAAGAAATAAACACTGACCTGATCTCTGAGAAATATTTACGCTCTTAAAAAGGG 1022
Db 641 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 660
QY 1023 GAAAGCCGTATGATCCCTTTGATCGCAGTCATCTGAAACTGTTCCACATCAATAGTT 1082

Db 661 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 680
QY 1083 GATGTCGATACCAACGAATTGCTTAAAGTAGCAGCTCTTAAACAGCTAGGACGTAAC 1142
Db 681 AspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 700
QY 1143 TTAGACTTCAGAGATTTATAGATCCTCGTGAAGGCTAAACTACTCTACAACAATCTC 1202
Db 701 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 720
QY 1203 GATCCTTTTGGTATTATGGACTATACCTTAACCTGAAAGTAGAGGATAATCACCATGAC 1262
Db 721 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 740
QY 1263 ACCAACCGTATCATCAACGGTTTATATGGGCAAGGACCCGAAAGAGAGAGATGTAGCTAT 1322
Db 741 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 760
QY 1323 CATTTAGCC 1331
Db 761 HisLeuAla 763

RESULT 13
AEB93425
ID AEB93425 standard; protein; 415 AA.
XX
AC AEB93425;
XX
DT 20-OCT-2005 (first entry)
XX
DE Streptococcus equisimilis streptokinase protein SEQ ID NO:1.
XX
KW protein quantitation; streptokinase.
XX
OS Streptococcus dysgalactiae subsp. equisimilis.
XX
PN US2005176085-A1.
XX
PD 11-AUG-2005.
XX
PF 15-NOV-2004; 2004US-00988943.
XX
PR 19-NOV-2003; 2003CU-00000269.
XX
PA (BETA/) BETANCOURT N L H.
PA (DORT/) DORTA-DUQUE J F D C.
PA (PERE/) PEREZ V A B.
PA (VALD/) VALDES J G.
PA (LOPE/) LOPEZ L J G.
PA (PALO/) PALOMARES G R P.
PA (FEYT/) FEYT R P.
PA (GILF/) GIL F M A.
XX
PI Betancourt NLH, Dorta-Duque JFD, Perez VAB, Valdes JG, Lopez LJG;
PI Palomares GRP, Feyt RP, Gil FMA;
XX
DR WPI; 2005-596348/61.
XX
PT Identifying and quantifying proteins in complex mixtures by selectively
PT isolating peptides not containing histidine nor arginine from each
PT protein, and determining the relative concentration of proteins in
PT different samples.
XX
PS Example 3; SEQ ID NO 1; 91pp; English.
XX
CC The invention relates to a method for identifying and quantifying one or
CC more proteins in complex mixtures by selectively isolating peptides not
CC containing histidine nor arginine (NHR peptides) from each protein, and
CC determining the relative concentration of one or more proteins in
CC different samples from the ratio between the areas of estinated
CC theoretical spectra for each NHR peptide labeled with different isotopes
CC in each sample. The method comprises: (a) enzymatically or chemically

CC hydrolyzing the sample or samples of proteins; (b) chemically modifying
 CC alpha and epsilon amino groups (alpha- and epsilon-NH₂) of every peptide
 CC obtained in step (a); (c) isolating the NHR peptides by cation exchange
 CC chromatography from the mixture of peptides obtained in step (b); (d)
 CC identifying proteins by mass spectrometry analysis of the NHR peptides
 CC obtained in step (c); (e) differential isotopic labeling of protein
 CC samples previously to step (a) or during steps (a) or (b) and immediately
 CC mixing at least a portion of the samples; and (f) relative quantifying of
 CC more proteins in the mixtures of step (e) from the ratio between the
 CC areas of estimated theoretical spectra of the pair of NHR peptides
 CC identified in step (d), as well as from the ratio between the areas of
 CC the estimated theoretical spectra of fragments from the NHR peptides,
 CC generated in step (d). Also described is a kit for the identification and
 CC quantification of proteins in complex mixtures, which comprises the
 CC method mentioned above. The method and kit are useful for identifying or
 CC quantifying proteins, such as those with vaccinal, therapeutic or
 CC diagnostic uses, in complex mixtures. The present sequence represents a
 CC streptokinase protein, which is used in an example from the present
 CC invention.

XX
 SQ Sequence 415 AA;

Alignment Scores:
 Pred. No.: 3,79e-176 Length: 415
 Score: 1963.00 Matches: 379
 Percent Similarity: 98.96% Conservative: 1
 Best Local Similarity: 98.70% Mismatches: 4
 Query Match: 66.79% Indels: 0
 DB: 9 Gaps: 0

US-09-940-235-10 (1-1661) x AEB93425 (1-415)

QY 180 ATGATGCTGGACCTGAGTGGCTGCTAGACCTGCTCATCTCTCAACACAGCCGATTTGGTT 239
 Db 1 MetilLeAlaGlyProGluThrLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
 QY 240 GTTACGCTTGTGGTACTGTTTCAGGGGAGCAATCAAGACATTAGTCTTAAATTTTGA 299
 Db 21 ValSerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPheGlu 40
 QY 300 ATCGATCTAACATCAGCCTCTCTCATGGAGGAAACAGACAGCAGCTTAAGTCCAAAA 359
 Db 41 IleAspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnLeuSerProLys 60
 QY 360 TCMAAACATTTGCTACTGATAGTGGCGGATGTCAATAAATTTGAGAAAGCTGACTTA 419
 Db 61 SerLysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGluLysAlaAspLeu 80
 QY 420 CTAAAGGCTATTCAAGCAATTTGATCGCTAGCTCCACAGTAAACGACGACTACTTTGAG 479
 Db 81 LeuLysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGlu 100
 QY 480 GTCAATTGATTTTGAAGCGATGCAACATTACTGATCGAAACGGCAAGCTCTACTTTGCT 539
 Db 101 VallieAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAla 120
 QY 540 GACAAAGATGTTTGGTAACTTGGCGACCCAACTGTCCAAAGATTTTGTAAACGCGA 599
 Db 121 AspLysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGly 140
 QY 600 CATGTGCGGTTAGACCATATAAGAAACCAATACAAACCAACGAGCAATCTGTGAT 659
 Db 141 HisValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAsp 160
 QY 660 GTGGAATATACGTGATGTTTACTCCCTTAAACCTGTGATGACGATTTTCAGACGAGTCTC 719
 Db 161 ValGluTyrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeu 180
 QY 720 AAAGATACCTAAGCTATTGAAACACCTAGCTATCGGTGACACCATCATCTCAAGAATTA 779
 Db 181 LysAspThrLysLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 200
 QY 780 CTAGCTCAAGCAACAGCATTTTAAACAAACCCAGGCTATACGATTTATGACGT 839

Db 201 LeuAlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArg 220
 QY 840 GACTCTCAATCTGCTCACTCATGACATGACATGATTTTCGTCAGATTTTACCAATGGATCAA 899
 Db 221 AspSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGln 240
 QY 900 GAGTTTACTTACCGTGTAAATAATCGGAACAAGCTTATAGGATCAATAAAAAATCTGGT 959
 Db 241 GluPheThrTyrHisValLysAsnArgGluGlnAlaTyrGluLeuAsnLysSerGly 260
 QY 960 CTGAATCAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAA 1019
 Db 261 LeuAsnGluGluLeuAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLys 280
 QY 1020 GGGGAAAGCCGATGATCCCTTTGATCGGAGTCACCTTGAACCTGTTTACCATCAATATAC 1079
 Db 281 GlyGluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyr 300
 QY 1080 GTTGATGTCATACCAACGAATTTGCTAAAGTGACGAGCTCTTAACAGCTAGCAGCT 1139
 Db 301 ValAspValAsnThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArg 320
 QY 1140 AACTTAGACTTCAGAGATTTATACGATCTCTGCTGATAAGGCTAAACTTCTTCAACAAT 1199
 Db 321 AsnLeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsn 340
 QY 1200 CTCGATGCTTTTGGTATTATGGACTATACCTTAACCTGGAAGAGTAGAGGATATACGAT 1259
 Db 341 LeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAsp 360
 QY 1260 GACACCAACCGTATCATACCGTTTATATGGCAAGGACCCGAGAGAGATGCTAGC 1319
 Db 361 AspThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSer 380
 QY 1320 TATCATTTAGCC 1331
 Db 381 TyrHisLeuAla 384
 RESULT 14
 AAW86143
 ID AAW86143 standard; protein; 414 AA.
 XX
 AC AAW86143;
 XX
 DT 17-OCT-2003 (revised)
 DT 03-MAR-1999 (first entry)
 XX
 DE Streptokinase (SK) protein sequence.
 XX
 KW Non-immunogenic; epitope; T-cell; immunogenicity; immune system; SK;
 immunogl bulin; therapeutic; streptokinase.
 OS Streptococcus dysgalactiae subsp. equisimilis.
 XX
 PN WO9852976-A1.
 XX
 PD 26-NOV-1998.
 XX
 PF 21-MAY-1998; 98WO-GB001473.
 XX
 PR 21-MAY-1997; 97GB-00010480.
 PR 31-JUL-1997; 97GB-00016197.
 PR 28-NOV-1997; 97GB-00025270.
 PR 02-DEC-1997; 97US-0067235P.
 PR 14-APR-1998; 98GB-00007751.
 XX
 PA (BIOV-) BIOVATION LTD.
 XX
 PI Carr FJ;
 XX
 DR WPI; 1999-045301/04.
 XX

PT Reducing immunogenicity of proteins - by modifying the amino acid
PT sequence of the protein to eliminate potential epitopes for T-cells of a
PT given species.

XX
XX
XX Example 6; Fig 28; 77pp; English.

CC The invention relates to a method for the production of non-immunogenic
CC proteins. The method comprises determining at least part of the amino
CC acid sequence of the protein; (b) identifying in the amino acid sequence
CC one or more potential epitopes for T-cells (T-cell epitopes) of the given
CC species; and (c) modifying the amino acid sequence to eliminate at least
CC one of the T-cell epitopes identified in step (b) thereby to eliminate or
CC reduce the immunogenicity of the protein when exposed to the immune
CC system of the given species. A method of analysing a pre-existing protein
CC to predict the basis for immunogenic responses is also provided. The
CC methods can be used particularly for reducing the immunogenicity of
CC immunoglobulins or therapeutic proteins, e.g. Streptokinase (SK). The
CC products can be used for diagnosis and therapy. The present sequence
CC represents the amino acid sequence of the SK protein. (Updated on 17-OCT-
CC 2003 to standardise OS field)

XX Sequence 414 AA;

Alignment Scores:

Pred. No.: 1.12e-175 Length: 414
Score: 1958.00 Matches: 378
Percent Similarity: 98.96% Conservative: 1
Best Local Similarity: 98.69% Mismatches: 4
Query Match: 66.62% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-10 (1-1661) x AAW86143 (1-414)

QY 183 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTCTCAACACCCCAATTGGTGT 242
Db 1 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
QY 243 AGCGTTGCTGGTACTGTTAGGCGACGAATCAAGACATTAGTCTTAAATTTTGAATC 302
Db 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspLeuSerLeuPheGluLeu 40
QY 303 GATCTAACATCAGCCTGCTCATGAGGAAAGACAGACGAAGCTTAAAGTCCAAATCA 362
Db 41 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 60
QY 363 AAACCAATTGCTACTGATAGTGGCGGATGTCATTAACCTTGAGAAAGCTGACTACTA 422
Db 61 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu 80
QY 423 AAGGCTATTCAAGAACAAATTGATCGCTTAACGCTCCACAGTAACGACGACTACTTTGAGTC 482
Db 81 LysAlaIleGlnGlnGluLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100
QY 483 ATTGATTTTGAAGCGATGCAACCATTTACTGATCGAAACGGCAGGCTACTTTGCTGAC 542
Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
QY 543 AAGATGGTTGGTAACTTGGCCGACCAACCTGTCGAAGAAATTTTGTGAAGCGACAT 602
Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
QY 603 GTGGCGTTAGACCATATAAGAAAACCAATATACAAAACCAAGCAATCTGTTGATGTG 662
Db 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
QY 663 GAATATACGTGTCAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCAAA 722
Db 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuLys 180
QY 723 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTA 782
Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 200

QY 783 GCTCAAGCACAAAGCATTTTAAACAAACACCCAGCTATACGATTATTAACAGTGCAC 842
Db 201 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp 220
QY 843 TCCTCAATCGTCACATCATGACAAATTTTCCGTACGATTTTACCAATGATCAAGAG 902
Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
QY 903 TTTACTTACCGTGTAAATAATCGGAAACAAGCTTATAGATCAATAAATAATCTGGCTCG 962
Db 241 PheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsnLysLysSerGlyLeu 260
QY 963 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAGCTCTTAAAAAGGG 1022
Db 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280
QY 1023 GAAAGCGGTATGATCCCTTTGATCGCAGTCATCTGAAACTGTTCACCAATACATGCTT 1082
Db 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
QY 1083 GATGTCGATACCAACGAATTGCTAAAGTGACGAGCTCTTAAACAGCTAGCGAACGTAAC 1142
Db 301 AspValAsnThrAsnGluLeuLeuLysSerGlnLeuLeuThrAlaSerGluArgAsn 320
QY 1143 TTAGACTTCAGAGATTTATACGATCCCTCGTATAGGCTAAACTACTCTACAAATCTC 1202
Db 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
QY 1203 GATCCTTTTGGTATTATGGACTATACCTTAACTCGAAAGTAGAGGATAATCACGATGAC 1262
Db 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
QY 1263 ACCAACCGTATCATACCGTTTATATGGCAGCAGCCCGAGAGAGAGAGATGCTAGCTAT 1322
Db 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 380
QY 1323 CATTTAGCC 1331
Db 381 HisLeuAla 383
RESULT 15
AAB01295
ID AAB01295 standard; protein; 414 AA.
XX AAB01295;
XX 12-SEP-2003 (revised)
DT 25-SEP-2000 (first entry)
XX
XX Wild type streptokinase from Streptococcus equisimilis.
DE
XX Immunogenicity; immunogen; T cell epitope; T-lymphocyte; drug; vaccine;
KW carrier; diagnosis; therapy; complement; C3; C5; CVF; plasminogen;
KW streptokinase; fibrin; blood clot; thrombolysis; plasmin;
KW myocardial infarction.
XX
OS Streptococcus dysgalactiae subsp. equisimilis.
XX
XX WO200034317-A2.
XX
XX 15-JUN-2000.
XX
XX 08-DEC-1999; 99WO-GB004119.
XX
XX 08-DEC-1998; 98GB-00026925.
XX
XX 02-FEB-1999; 99GB-00002139.
XX
XX (BIOV-) BIOVATION LTD.
XX
XX Carr FJ, Adair FS, Hamilton AA, Carter G;
XX
XX WPI; 2000-423372/36.
XX

PT Rendering a protein non-immunogenic or less immunogenic useful in
 PT medicine and in diagnostics involves determining the amino acid sequence
 PT of the protein, identifying and modifying potential epitopes.

XX Example 2; Fig 5; 42pp; English.

CC Proteins or their fragments can be rendered non-immunogenic or less
 CC immunogenic by identifying one or more potential T cell epitopes and
 CC modifying the sequence to eliminate at least one of these T cell epitopes
 CC to reduce the immunogenicity of the protein when exposed to the immune
 CC system of another organism. The method can be used for producing a
 CC protein with an enzymatic activity which has a beneficial therapeutic
 CC effect, a protein used to convert inactive drugs to its active form
 CC within a living organism, a vaccine, a protein used as a carrier of other
 CC molecule or a protein which binds to other molecules within or introduced
 CC within the living organism in order to alter the bio distribution of
 CC other molecules, such that the protein does not stimulate immune response
 CC in the living organism. The less immunogenic protein is useful in
 CC medicine, diagnosis and in manufacture of a therapeutic or a diagnostic
 CC agent. Streptokinase is produced by certain strains of beta-haemolytic
 CC streptococci. The protein has no inherent enzymatic activity but has
 CC considerable clinical importance owing to its ability to efficiently bind
 CC human plasminogen, potentiating its activation to plasmin and thereby
 CC promoting the dissolution of fibrin filaments in blood clots.
 CC Streptokinase is an effective thrombolytic agent in the treatment of
 CC coronary thrombosis, improving survival and preserving left ventricular
 CC function following myocardial infarction. The native protein is
 CC immunogenic and the production of neutralising antibodies in humans
 CC generally limits the protein to a single use. The new method could
 CC provide a longer therapeutic use for streptokinase. This is the sequence
 CC of the wild type streptokinase. The altered streptokinase sequence is
 CC given in AAB01296. See GENESQ records AAB01289-B01302. (Updated on 12-
 CC SEP-2003 to standardise OS field)

XX Sequence 414 AA;

Alignment Scores:
 Pred. No.: 1,12e-175 Length: 414
 Score: 1958.00 Matches: 378
 Percent Similarity: 98.96% Conservative: 1
 Best Local Similarity: 98.69% Mismatches: 4
 Query Match: 66.62% Indels: 0
 Ds: 3 Gaps: 0

US-09-940-235-10 (1-1661) x AAB01295 (1-414)

QY 183 ATGTGTCGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAAATGGTGT 242
 Db 1 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal 20
 QY 243 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC 302
 Db 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 40
 QY 303 GATCTAACATCAGCACCTCTCATGAGGAGACAGACAGCCAGGCTTAAGTCCAAATCA 362
 Db 41 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 60
 QY 363 AAACCATTTGCTACTGATAGTGGCGGATGTACATATAAACTTGAGAAAGCTGACTTACTA 422
 Db 61 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu 80
 QY 423 AAGGCTATTTCAGAACAAATGATCGCTAACGTCACAGTAACGACGACTACTTTGAGGTC 482
 Db 81 LysAlaIleGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100
 QY 483 ATTGATTTTTCAGCGCATCAACATTACTGATCGAAACGGCAAGGCTACTTCTTGGCTGAC 542
 Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
 QY 543 AAAGATGGTTCCGTAACTTGGCGACCCAACTGTCCAAAGAAATTTTGTAAAGCGGACAT 602
 Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140

QY 603 GTGCGCGTTAGACCATATAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 662
 Db 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
 QY 663 GAATATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTCAAA 722
 Db 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
 QY 723 GATCTAAGCTATTGAAAAACACATAGCTATCGGTGACACCATCATCTCAAGAAATTTACTA 782
 Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 200
 QY 783 GCTCAGACCAAGAGCATTTTAAACCAAAACCCAGGCTATACGATTTTATGACGCTGAC 842
 Db 201 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp 220
 QY 843 TCCTCAATCTCTACTCATGACATGACATTTTCGTCAGCATTTTACCAATGGATCAAGAG 902
 Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
 QY 903 TTTACTTACCGTGTAAAAATCGGAAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 962
 Db 241 PheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsnLysLysSerGlyLeu 260
 QY 963 AATCAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1022
 Db 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280
 QY 1023 GAAAGCCGATGATCCCTTTGATCGCAGTCACCTGAAACTGTTCAACATCAAAATACGTT 1082
 Db 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
 QY 1083 GATGTCGATACCAACGAATTTGCTTAAAGTGAAGCAGCTCTTAAACAGCTAGCGAACGTAAC 1142
 Db 301 AspValAsnThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320
 QY 1143 TTAGACTTCAGAGATTTTATACGATCTCTCGTGATAGGCTAAACTACTCTACACAAATCTC 1202
 Db 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
 QY 1203 GATGCTTTTGGTATTATGAGCTATACCTTAACCTGGAAGAGTAGAGGATAATCAGATGAC 1262
 Db 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
 QY 1263 ACCAACCGTATCATACCCGTTTATATGGCAAGCCGACCCGAGAGAGAGAAATGCTAGCTAT 1322
 Db 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 380
 QY 1323 CATTTAGCC 1331
 Db 381 HisLeuAla 383

Search completed: January 28, 2006, 02:08:17
 Job time : 132.543 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 28, 2006, 01:49:55 ; Search time 23.6951 Seconds
(without alignments)
11590.966 Million cell updates/sec

Title: US-09-940-235-10
Perfect score: 2939
Sequence: 1 gcaacccgcagcctagcc.....gaataagctgtaccatctaa 1661

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0
Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 1144120

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_n2p.model -DEV=xlp
-Q=/Cgn2_1/USFTO_spool_p/US09940235/runat_27012006_144219_27602/app_query.fasta_1.7708
-DB=Issued_Patents_AA -OFT=fastan -SUFFIX=rai -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09940235 @Cgn_1_159 @runat_27012006_144219_27602 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_AA.*

- 1: /Cgn2_6/prodata/1/iaa/5/COMB.pep.*
- 2: /Cgn2_6/prodata/1/iaa/6/COMB.pep.*
- 3: /Cgn2_6/prodata/1/iaa/H/COMB.pep.*
- 4: /Cgn2_6/prodata/1/iaa/PCTUS/COMB.pep.*
- 5: /Cgn2_6/prodata/1/iaa/RE/COMB.pep.*
- 6: /Cgn2_6/prodata/1/iaa/backfiles/pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1984	67.5	413	2	US-10-360-101-264
2	1984	67.5	414	2	US-09-211-542A-6
3	1984	67.5	414	6	5240845-1
4	1984	67.5	440	1	US-08-560-098A-52
5	1984	67.5	795	2	US-09-211-542A-2
6	1966.5	66.9	813	1	US-08-488-940-3
7	1966.5	66.9	1194	1	US-08-488-940-1
8	1963	66.8	415	1	US-07-854-596B-26
9	1958	66.6	435	1	US-07-854-596B-19
10	1958	66.6	440	1	US-07-854-596B-15
11	1958	66.6	483	1	US-07-854-596B-43
12	1958	66.6	483	1	US-07-854-596B-47

13	1958	66.6	499	1	US-07-854-596B-28	Sequence 28, Appl
14	1958	66.6	859	1	US-07-854-596B-35	Sequence 35, Appl
15	1936.5	65.9	1194	1	US-08-488-940-17	Sequence 17, Appl
16	1936.5	65.9	1194	1	US-08-488-940-18	Sequence 18, Appl
17	1928	65.6	413	1	US-08-759-559-12	Sequence 12, Appl
18	1928	65.6	413	2	US-09-294-457-12	Sequence 12, Appl
19	1928	65.6	413	2	US-09-919-703-12	Sequence 12, Appl
20	1902	64.7	384	2	US-09-374-038-4	Sequence 4, Appl
21	1902	64.7	384	2	US-09-658-179-4	Sequence 4, Appl
22	1900	64.6	800	1	US-08-488-940-4	Sequence 4, Appl
23	1900	64.6	1181	1	US-08-488-940-2	Sequence 2, Appl
24	1898	64.6	372	2	US-09-374-038-3	Sequence 3, Appl
25	1898	64.6	372	2	US-09-658-179-3	Sequence 3, Appl
26	1887	64.2	401	2	US-09-374-038-1	Sequence 1, Appl
27	1887	64.2	401	2	US-09-658-179-1	Sequence 1, Appl
28	1887	64.2	413	2	US-09-374-038-2	Sequence 2, Appl
29	1887	64.2	413	2	US-09-658-179-2	Sequence 2, Appl
30	1879	63.9	747	1	US-07-854-596B-40	Sequence 40, Appl
31	1875	63.8	369	1	US-07-854-596B-31	Sequence 31, Appl
32	1695.5	57.7	736	2	US-09-211-542A-4	Sequence 4, Appl
33	1686	57.4	356	2	US-09-211-542A-12	Sequence 12, Appl
34	785	26.7	150	2	US-09-211-542A-14	Sequence 14, Appl
35	600	20.4	2231	1	US-08-153-799-16	Sequence 16, Appl
36	600	20.4	2324	1	US-08-283-857-1	Sequence 1, Appl
37	600	20.4	2324	4	PCT-US95-09819-1	Sequence 1, Appl
38	600	20.4	2327	6	5455158-1	Patent No. 5455158
39	600	20.4	2386	1	US-09-016-366A-12	Sequence 12, Appl
40	600	20.4	2446	1	US-08-551-356-2	Sequence 2, Appl
41	600	20.4	2446	4	PCT-US93-12687-2	Sequence 2, Appl
42	596	20.3	2355	2	US-10-360-101-235	Sequence 235, App
43	596	20.3	2386	2	US-09-961-403-1	Sequence 1, Appl
44	283	9.6	286	2	US-09-078-091-4	Sequence 4, Appl
45	275	9.4	286	2	US-09-078-091-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-10-360-101-264
; Sequence 264, Application US/10360101
; Patent No. 6861236
; GENERAL INFORMATION:
; APPLICANT: Moll, Gert N.
; APPLICANT: Leenhouts, Cornelis J.
; TITLE OF INVENTION: Export and modification of (poly)peptide in the lantibiotic way
; FILE REFERENCE: 2183-5673
; CURRENT APPLICATION NUMBER: US/10/360,101
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: EP 02077060.8
; PRIOR FILING DATE: 2002-05-24
; NUMBER OF SEQ ID NOS: 309
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 264
; LENGTH: 413
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: sequence of streptokinase
US-10-360-101-264

Alignment Scores:
Pred. No.: 8.87e-186 Length: 413
Score: 1984.00 Matches: 383
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 67.51% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-10 (1-1661) x US-10-360-101-264 (1-413)

Qy 183 ATTGCTGGACCTGAGTGGCTGCTAGACGCTCCATCTGTCACACACGCAATGGTGTGT 242
Db 1 ILEAlaGlyProGluIrrPLeuAenAenSerGlnLeuValVal 20

QY 243 AGCGTTGCTGCTAGTCTGAGGGACGAATCAAGACATTAGTCTTAAATTTTTTTGAAATC 302
 Db |||||||
 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 40
 QY 303 GATCTAACATCAGCACTGCTCATGGAGGAAGACAGACAGCGTTAAAGTCCAAATCA 362
 Db |||||||
 41 AspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 60
 QY 363 AAACCAATTTGCTAGTGTGGCGGATGTCATAAACTTGAGAAAGCTGACTACTA 422
 Db |||||||
 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 80
 QY 423 AAGGCTATTCAAGAACAAATTGATCGCTAACGTCACAGTAACAGCACTACTTTGAGGTC 482
 Db |||||||
 81 LysAlaIleGlnGluLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100
 QY 483 ATTGATTTTGCAGGATCAACCACTTACTGATCGAAACGGCAAGCTCTACTTTGCTGAC 542
 Db |||||||
 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
 QY 543 AAAGATGGTTGGTAACCTGCGGACCCAACTGTCCTCAAGAAATTTTGTCTAAGCGGACAT 602
 Db |||||||
 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
 QY 603 GTGCGCTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTTTGATGTG 662
 Db |||||||
 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
 QY 663 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGCTCTCAAA 722
 Db |||||||
 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuLys 180
 QY 723 GATCTAAGCTATTGAAACACTAGTATCGGTGACACCATCAATCTCAAGAAATTTACTA 782
 Db |||||||
 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 200
 QY 783 GCTGAAGCAGCAAGCATTTTAAACAAACCCAGGCTATACGATTTTATGACGTCGAC 842
 Db |||||||
 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 220
 QY 843 TCCTCAATCGTCACATGACATGACATTTTCGTCGACGATTTTACCAATGGATCAAGAG 902
 Db |||||||
 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
 QY 903 TTTACTTACCGTGTAAAAATCGGAACAAGCTTTATPAGGATCAATAAAAAATCTGCTCTG 962
 Db |||||||
 241 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 260
 QY 963 AATGAAGAAATAACCACTGACCTGATCTCTGAGAAATATTAAGTCTTAAAAAAGG 1022
 Db |||||||
 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280
 QY 1023 GAAAAGCCGTATGATCCCTTTGATCGCTGACGCTGAAACTGTTCAACATCAATACGTT 1082
 Db |||||||
 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
 QY 1083 GATGTCGATACCAACGAATTTGCTTAAAGTGAAGCTCTTAAACAGCTAGCGCAACGTAAC 1142
 Db |||||||
 301 AspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320
 QY 1143 TTAGACTTCAGAGATTTATACATCTCTGTGATAAGGCTAAACTACTCTACAAATCTC 1202
 Db |||||||
 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
 QY 1203 GATGCTTTTGTATTATGACTATACCTTAACTGGAAGTAGAGGATATACCATGAC 1262
 Db |||||||
 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
 QY 1263 ACCAACCGTATCATACCGTTTATATGCGCAAGCGACCCGAAAGGAGAGAAATGCTAGCTAT 1322
 Db |||||||
 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 380

QY 1323 CATTAGCC 1331
 Db |||||||
 381 HisLeuAla 383
 RESULT 2
 US-09-211-542A-6
 : Sequence 6, Application US/09211542A
 : Patent No. 6210667
 : GENERAL INFORMATION:
 : APPLICANT: Reed, Guy L.
 : TITLE OF INVENTION: BACTERIAL FIBRIN-DEPENDENT PLASMINOGEN ACTIVATOR
 : NUMBER OF SEQUENCES: 14
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: BROMBERG & SUNSTEIN, LLP
 : STREET: 125 Summer Street
 : CITY: Boston
 : STATE: Massachusetts
 : COUNTRY: USA
 : ZIP: 02110
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : OPERATING SYSTEM: IBM PC compatible
 : SOFTWARE: Patent In Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/09/211,542A
 : FILING DATE: 15-December-1998
 : CLASSIFICATION: 1653
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 60/069,497
 : FILING DATE: 15-December-1997
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Attorney, Strimpel, Harriet M.
 : REGISTRATION NUMBER: 37,008
 : REFERENCE/DOCKET NUMBER: 1874/111
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (617) 443-9292
 : TELEFAX: (617) 443-0004
 : INFORMATION FOR SEQ ID NO: 6:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 414 amino acids
 : TYPE: amino acid
 : TOPOLOGY: linear
 : MOLECULE TYPE: protein
 US-09-211-542A-6
 Alignment Scores:
 Pred. No.: 8 88e-186 Length: 414
 Score: 1984.00 Matches: 383
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 67.51% Indels: 0
 DB: 2 Gaps: 0
 US-09-940-235-10 (1-1661) x US-09-211-542A-6 (1-414)
 QY 183 ATTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAAATGGTTGTT 242
 Db |||||||
 1 IleAlaGlyProGluTyrLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
 QY 243 AGCGTTGCTGCTAGTGTGGGGACGAATCAAGACATTAAGTCTTAAATTTTTTGAATC 302
 Db |||||||
 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 40
 QY 303 GATCTAACATCAGCACTGCTCATGGAGGAAGACAGACAGCGTTAAAGTCCAAATCA 362
 Db |||||||
 41 AspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 60
 QY 363 AAACCAATTTGCTACTGATGTCGCGGATGTCATAAACTTGAGAAAGCTGACTACTA 422
 Db |||||||
 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 80
 QY 423 AAGGCTATTCAAGAACAAATTTGATCGCTAACGTCACAGTAACAGCACTACTTTGAGGTC 482

```
Db 81 LysAlaIleGlnGluLeuIleAlaAsnValHisSerAsnAspTyrPheGluVal 100
QY 483 ATTGATTTCGAAAGCGATGCAACCATCTACTGATCGAAACCGCAAGGCTCTACTTTGCTGAC 542
Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
QY 543 AAGATGTTTCGGTAACTTCGCGCAACCAACCTGTCTCAAGAAATTTTGTAAAGCGGACAT 602
Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnPheLeuLeuSerGlyHis 140
QY 603 GTGCGGTAGACCATATAAGAAACCAACCAATACAAACCAACCAAGCGAAATCTGTTGATGTG 662
Db 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
QY 663 GAATATCTACTGACAGTTTACTCCCTTAAACCCCTGATCAGCATTTTCAGACAGGCTCAAA 722
Db 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
QY 723 GATCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTTACTA 782
Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 200
QY 783 GCTCAAGCAACAGCAATTTAAACAAACCAACCAACCGGCTATACGATTATGAAACGTGAC 842
Db 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 220
QY 843 TCTCTCAATCGTCTCATGACATGACATTTTCGTACGATTTTACCAATGATCAAGAG 902
Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
QY 903 TTTACTACCGTGTAAAATCGGGAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 962
Db 241 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 260
QY 963 AATGAAGAAATAACAACTGACCTGATCTCTGAGAAATATTACGCTCCTTAAAAAAGGG 1022
Db 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280
QY 1023 GAAAGCCGTATGATCCCTTTGATCGCAGTCACCTTGAACCTTGAACCTTCAACATACGTT 1082
Db 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
QY 1083 GATGTCGATACCAACGAATTCCTTAAAGAGTACGAGCTCTTAAACAGCTAGCGACGTAAC 1142
Db 301 AspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320
QY 1143 TTAGACTTCAGAGATTATACGATCCCTGATAGCTTAAAGCTTAACTCTTCAACAATCTC 1202
Db 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
QY 1203 GATGCTTTTGGTATTATGACTATACCTTAACCTGGAAAGTAGAGGATAATACAGATGAC 1262
Db 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
QY 1263 ACCAACCGTATCATACCGTTTATATAGGCAACCGGACCGCAAGAGAGATGCTAGCTAT 1322
Db 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 380
QY 1323 CATTTAGGCC 1331
Db 381 HisLeuAla 383
```

RESULT 3

5240845-1

; Patent No. 5240845

; APPLICANT: FUJII, SETSURO; TAKADA, KAORUKO.; KATANO, TAMIKI;

; MAJIMA, EIJI; OGINO, KOICHI; ONO, KENJI; SAKATA, YASUYO; UENOYAMA,

; TSUTOMU

; TITLE OF INVENTION: MUTATED STREPTOKINASE PROTEINS

; NUMBER OF SEQUENCES: 65

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/549,049

; FILING DATE: 06-JUL-1990

; SEQ ID NO:1:

; LENGTH: 414

5240845-1

Alignment Scores:

Pred. No.:	8.88e-186	Length:	414
Score:	1984.00	Matches:	383
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	67.51%	Indels:	0
DB:	6	Gaps:	0

US-09-940-235-10 (1-1661) x 5240845-1 (1-414)

```
QY 183 ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCAATGGTGTGTT 242
Db 1 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
QY 243 AGCGTTGCTGCTACTGTTGAGGGACGAATCAAGACATTAGTCTTTAAATTTTTTCAAATC 302
Db 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 40
QY 303 GATCTAATCATCAGCACCTGCTCATGGAGGAAAGACAGACGAGGCTTAAGTCCAAATCA 362
Db 41 AspLeuThrSerArgProAlaHisGlyGlyLysThrGlnGlnGlyLeuSerProLysSer 60
QY 363 AAACCATTTGCTACTGATAGTGGCGGATGTACATAAACTTGAGAAAGCTGACTTACTA 422
Db 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeu 80
QY 423 AAGCTATTCAAGAACCAATTGATCGCTAAACGTCCACAGTACACGACGACTCTTTGAGTGC 482
Db 81 LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100
QY 483 ATTGATTTTGCACGCGATGCAACCATTTACTGATCGAAACGGCAAGGCTCTACTTTGCTGAC 542
Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
QY 543 AAGATGTTTCGGTAACTTCGCGCAACCAACCTGTCTCAAGAAATTTTGTCAAGCGGACAT 602
Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnPheLeuLeuSerGlyHis 140
QY 603 GTGCGCTTAGACCATATAAGAAACCAACCAATACAAACCAACCGAATCTGTTGATGTG 662
Db 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
QY 663 GAATATCTACTGATACGATTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCAAA 722
Db 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
QY 723 GATCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTTACTA 782
Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 200
QY 783 GCTCAAGCAACAGCAATTTTAAACAAACCAACCAACCGGCTATACGATTATGAAACGTGAC 842
Db 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 220
QY 843 TCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCATGATGATCAAGAG 902
Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
QY 903 TTTACTTACCGTGTAAAATCGGGAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 962
Db 241 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 260
QY 963 AATGAAGAAATAACAACTGACCTGATCTCTGAGAAATATTACGCTCCTTAAAAAAGGG 1022
Db 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280
QY 1023 GAAAGCCGTATGATCCCTTTGATCGCAGTCACCTTGAACCTTGAACCTTCAACATACGTT 1082
```


Db 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
 QY 1083 GATGTCGATACCAAGAAATGCTTAAAGCTGAGCAGCTCTTAAACAGCTAGCGAAGCTAAC 1142
 Db 301 AspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320
 QY 1143 TTAGACTTCAGAGATTATACGATCCTCGTGATAAGGCTAAACTACTCTACAAATCTC 1202
 Db 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
 QY 1203 GATGCTTTTGGTATTATGCACTACTTAACTTAACTGGAAGATAGAGGATATCAGATGAC 1262
 Db 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
 QY 1263 ACCAACCGTATCATACCGTTTATATGGCGAAGCGACCGAAGGAGAGAGATGCTAGCTAT 1322
 Db 361 ThrAsnArgIleIleThrValIleThrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 380
 QY 1323 CATTTAGGCC 1331
 Db 381 HisLeuAla 383

RESULT 4

US-08-560-098A-52
 ; Sequence 52, Application US/08560098A
 ; Patent No. 5976841
 ; GENERAL INFORMATION:
 ; APPLICANT: WNEUET, Stephan
 ; APPLICANT: HEINZEL-WIELAND, Regina
 ; APPLICANT: STEFFENS, Gerd Josef
 ; TITLE OF INVENTION: Proteins having Fibrinolytic and
 ; TITLE OF INVENTION: Coagulation-inhibiting Properties
 ; NUMBER OF SEQUENCES: 60
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: Evenson, McKeown, Edwards & Lenahan
 ; STREET: 1200 G Street, N.W., Suite 700
 ; CITY: Washington
 ; STATE: DC
 ; COUNTRY: USA
 ; ZIP: 20005
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/560,098A
 ; FILING DATE: 17-NOV-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: P 44 40 892.7
 ; FILING DATE: 17-NOV-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: EVANS, Joseph D.
 ; REGISTRATION NUMBER: 26,269
 ; REFERENCE/DOCKET NUMBER: 148/42448
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 628-8800
 ; TELEFAX: (202) 628-8844
 ; INFORMATION FOR SEQ ID NO: 52:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 440 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-560-098A-52
 Alignment Scores:
 Pred. No.: 9,098-186 Length: 440
 Score: 1984.00 Matches: 383
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 67.51% Indels: 0

DB: 1 Gaps: 0
 US-09-940-235-10 (1-1661) x US-08-560-098A-52 (1-440)
 QY 183 ATTGCTGGAGCTGAGTGGCTGCTAGACCGTCCATCTCTCAACACAGCCCAATGGTGTGTT 242
 Db 27 IleAlaGlyProGluTyrLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal 46
 QY 243 AGCGTTGCTGCTACTGTTGAGGGGCGAATCAAGACATTAGTCTTAAATTTTGAATC 302
 Db 47 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 66
 QY 303 GATCTAACTACAGACCTGCTCATGGAGGAAAGACAGACGAAAGCTTAAAGTCCAAAAATCA 362
 Db 67 AspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 86
 QY 363 AAACCATTTGCTACTGATAGTGGCGGATGTCATAAATCTTGAGAAAGCTGACTTACTA 422
 Db 87 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 106
 QY 423 AAGGCTATTCAAGACCAATTTGATCGCTTAACGTCCACAGTAACGACGACTACTTTGAGTTC 482
 Db 107 LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 126
 QY 483 ATTGATTTTGCAGAGCGATGCAACCATTTACTGTGATCGAAACCGCAAGGTCTACTTTGCTGAC 542
 Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146
 QY 543 AAAGATGGTTCGTAACCTTTGCGACCCCAACCTGTGTCCAGAAATTTTGTAAAGCGGACAT 602
 Db 147 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 166
 QY 603 GTGCGGTAGACCATATAAAGAAAACCAATACAAACCAAGCGAAATCTGTGATGTCG 662
 Db 167 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 186
 QY 663 GAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACCATTTTCAGACAGGCTCTCAA 722
 Db 187 GluTyrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuLys 206
 QY 723 GATPACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAATTTACTA 782
 Db 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 226
 QY 783 GCTCAAGCACAAAGCAATTTTAAACAAACCAACCCAGGCTATAGGATTTATGAACGTGAC 842
 Db 227 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 246
 QY 843 TCCTCAATCGTCACTCATGACAAATGACATTTTCGTACGATTTTACCAATGGATCAAGAG 902
 Db 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
 QY 903 TTTTACTTACCGTGTAAAAATCGGGAACCAAGCTTTATAGGATCAATAAAAAATCTGTCGTG 962
 Db 267 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 286
 QY 963 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCTCTTAAAAAGGG 1022
 Db 287 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 306
 QY 1023 GAAAGCCGTATGATCCCTTTGATCGCAGTCACCTGAAACCTGTTCCACCATCAATACGTT 1082
 Db 307 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326
 QY 1083 GATGTCGATACCAACGAATTTGCTAAAGGTGAGCAGCTCTTAAACAGCTAGCGAAGCTAAC 1142
 Db 327 AspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
 QY 1143 TTAGACTTCAGAGATTATACGATCCTCGTGATAAGGCTAAACTACTCTCAACAATCTC 1202
 Db 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
 QY 1203 GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAGATAGAGGATATACGAGTAC 1262

Db 367 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 386
QY 1263 ACCAACCGTATCAACCGTTTATATGGCAAGCGACCCGGAAGGAGAGAAATGCTAGCTAT 1322
Db 387 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 406
QY 1323 CATTTAGCC 1331
Db 407 HisLeuAla 409

RESULT 5

US-09-211-542A-2
; Sequence 2, Application US/09211542A
; Patent No. 6210667
; GENERAL INFORMATION:
; APPLICANT: Reed, Guy L.
; TITLE OF INVENTION: BACTERIAL FIBRIN-DEPENDENT PLASMINOGEN ACTIVATOR

; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROMBERG & SUNSTEIN, LLP
; STREET: 125 Summer Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/211,542A
; FILING DATE: 15-December-1998

; CLASSIFICATION: 1653
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/069,497
; FILING DATE: 15-December-1997

; ATTORNEY/AGENT INFORMATION:
; NAME: Attorney, Strimpel, Harriet M.
; REGISTRATION NUMBER: 37,008
; REFERENCE/DOCKET NUMBER: 1874/111

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 443-9292
; TELEFAX: (617) 443-0004

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 795 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-09-211-542A-2

Alignment Scores:
Pred. No.: 1,15e-185 Length: 795
Score: 1984.00 Matches: 383
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 67.51% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-10 (1-1661) x US-09-211-542A-2 (1-795)

QY 183 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACGCCCAATTTGGTTGT 242
Db 382 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal 401
QY 243 ACCGTGCTGCTACTGTTGAGGGGAGCAATCAAGACATTAAGTCTTAAATTTTGAATC 302
Db 402 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 421
QY 303 GATCTAAACATCAGACCTGCTCATGGGAAAGACAGACGAGCTTAAGTCCAAATCA 362

Db 422 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 441
QY 363 AAACCATTTGCTACTGATAGTGGCCGCGATGTACACATAAACTTGAGAAAGCTGACTACTA 422
Db 442 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 461
QY 423 AAGGCTATTCAAGAACAAATTGATCCGCTAACACGTCCACAGTAAACGACGACTACTTTTCAGGTC 482
Db 462 LysAlaIleGlnGluLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 481
QY 483 ATTGATTTTGCAGCGATGCAACCAATTACTGATCGAAACGGCAAGGTCTACTTTCTGAC 542
Db 482 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 501
QY 543 AAAGATGGTTCGGTAAACCTTCCCGACCCCAACTGTGTCGAAGATTTTTCGTAAGCGGACAT 602
Db 502 LysAspGlySerValThrLeuProThrGlnProValGlnPheLeuLeuSerGlyHis 521
QY 603 GTGCGCGTTAGACCATATAAAGAAAAACAATACAAAAACCAAGCCGAAATCTGTTGATGTG 662
Db 522 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 541
QY 663 GAATATCTGTACAGATTACTCCCTTAAACCTGATGACGATTTTCAGACCGAGTCTCAAA 722
Db 542 GluTyrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuLys 561
QY 723 GATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCACATCTCAGAATTACTA 782
Db 562 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 581
QY 783 GCTCAAGCACAAAGCATTTTAAACCAAAACCAACCCAGGCTATACGATTTATGACGTCAC 842
Db 582 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 601
QY 843 TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTTACCAATGGATCAACAG 902
Db 602 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 621
QY 903 TTTACTTACCGTGTATAAAATTCGGGAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 962
Db 622 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 641
QY 963 AATCAAGAAATAAACAACACCTGATCTCTGAGAAATATAGTCTCTTAAAAAAGG 1022
Db 642 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysGly 661
QY 1023 GAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCCACCATCAAAATACGT 1082
Db 662 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 681
QY 1083 GATCTCGATACCAACGAATTGCTAAAAAGTGAGCAGCTCTTTAAGAGCTAGCGAAGCTAAC 1142
Db 682 AspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 701
QY 1143 TTAGACTTCAGAGATTTTATACGATCCCTCGATAGGCTAAACTACTCTACACAAATCTC 1202
Db 702 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 721
QY 1203 GATCTTTTGTATTATGACTATACCTTAACCTGAGAAAGTAGAGGATAATCAGGATCAC 1262
Db 722 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 741
QY 1263 ACCAACCGTATCAACCGTTTATATGGCAAGCGACCCGGAAGGAGAGAAATGCTAGCTAT 1322
Db 742 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 761
QY 1323 CATTTAGCC 1331
Db 762 HisLeuAla 764

RESULT 6

US-08-488-940-3

; Sequence 3, Application US/08488940

Patent No. 5854049
 GENERAL INFORMATION:
 APPLICANT: Reed, Guy L.
 TITLE OF INVENTION: PLASMIN-RESISTANT STREPTOKINASE
 NUMBER OF SEQUENCES: 20
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson P.C.
 STREET: 225 Franklin Street
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02110-2804
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/488,940
 FILING DATE: 09-JUN-1995
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: Fraser, Janis K.
 REGISTRATION NUMBER: 34,819
 REFERENCE/DOCKET NUMBER: 05433/009001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617/542-5070
 TELEFAX: 617/542-8906
 TELEX: 200154
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 813 amino acids
 TYPE: amino acid
 STRANDEDNESS: not relevant
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-488-940-3

Alignment Scores:
 Pred. No.: 6, 02e-184 Length: 813
 Score: 1966.50 Matches: 382
 Percent Similarity: 99.74% Conservative: 0
 Best Local Similarity: 99.74% Mismatches: 0
 Query Match: 66.91% Indels: 1
 DB: 1 Gaps: 1
 US-09-940-235-10 (1-1661) x US-08-488-940-3 (1-813)

QY 183 ATTGCTGACCTGAGTGGCTGTAGACCTGTCATCTGTCAACACAGCCAAATGGTTGTT 242
 DB 1 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
 QY 243 AGCGTTGCTGTGTTGTTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTGTGAATC 302
 DB 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPheGluIle 40
 QY 303 GATCTAACATCAAGACCTGCTCATGGAGGAGAAAGACAGACAGGCTTAAGTCCAAATCA 362
 DB 41 AspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 60
 QY 363 AAACCATTTGCTGATGATGTCGCGGATGCACATAAACTTCAGAAAGCTGACTTACTA 422
 DB 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlnLysAlaAspLeu 80
 QY 423 AAGGCTATTCAAGAACAAATTTGCTGCTTAACGTCACAGTAAACGACGACTACTTTGAGTTC 482
 DB 81 LysAlaIleGlnGluLeuIleAlaAsnValHisSerAsnAspAspThrPheGluVal 100
 QY 483 ATTGATTTTGCAGCGATGCAACCAATTAAGTATCGAAACGGCAAGTCTACTTTGCTGAC 542
 DB 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValThrPheAlaAsp 120
 QY 543 AAAGATGTTTCGTTAACTTTCGCGACCCCAACCTGTCTCAAGAAATTTTGTCTAAGCGGACAT 602

DB 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
 QY 603 GTGCGGTTAGACCATATAAGAAACCAATCAACCAACCAAGCGAATCTGTTGATGTCG 662
 DB 141 ValArgValArg--TyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 159
 QY 663 GAATATACTCTACAGATTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAA 722
 DB 160 GluTyrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuLys 179
 QY 723 GATACTAAGCTATTGAAACACACTAGCTATCGGTGACACCATCAATCTCAAGAATTA 782
 DB 180 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 199
 QY 783 GCTCAAGCACAAAGCATTTTAAACCAAAACCCAGGCTATAGATTATGACGTCAC 842
 DB 200 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 219
 QY 843 TCCTCAATCTCACTCATGCAATGACATTTTCGTACGATTTTACCAATGGATCAAGAG 902
 DB 220 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 239
 QY 903 TTTTACTTACCGTGTATAAAATCGGGAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 962
 DB 240 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 259
 QY 963 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1022
 DB 260 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysGly 279
 QY 1023 GAAAAGCGTATGATCCCTTTGATCGCAGTCACCTTGAACCTGTTCCACATCAAAATAGCTT 1082
 DB 280 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 299
 QY 1083 GATGTCGATACCAACCAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 1142
 DB 300 AspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 319
 QY 1143 TTAGACTTCAGAGATTTATACGATCTCTGCTGTAAGGCTAAACTCTCTCAACAATCTC 1202
 DB 320 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 339
 QY 1203 GATGCTTTGTTGTTATGAGTATACCTTAACTTGAAGAAAGTACGAGGATATCAGCATGAC 1262
 DB 340 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 359
 QY 1263 ACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAAAGAGAGAATGCTAGCTAT 1322
 DB 360 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 379
 QY 1323 CATTTAGCC 1331
 DB 380 HisLeuAla 382
 RESULT 7
 US-08-488-940-1
 Sequence 1, Application US/08488940
 Patent No. 5854049
 GENERAL INFORMATION:
 APPLICANT: Reed, Guy L.
 TITLE OF INVENTION: PLASMIN-RESISTANT STREPTOKINASE
 NUMBER OF SEQUENCES: 20
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson P.C.
 STREET: 225 Franklin Street
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02110-2804
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/488,940
 ; FILING DATE: 09-JUN-1995
 ; CLASSIFICATION: 514
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Fraser, Janis K.
 ; REGISTRATION NUMBER: 34,819
 ; REFERENCE/DOCKET NUMBER: 05433/009001
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 617/542-5070
 ; TELEFAX: 617/542-8906
 ; TELEX: 200154
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1194 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: not relevant
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-488-940-1

Alignment Scores:
 Pred. No.: 7,01e-184 Length: 1194
 Score: 1966.50 Matches: 382
 Percent Similarity: 99.74% Conservative: 0
 Best Local Similarity: 99.74% Mismatches: 0
 Query Match: 66.91% Indels: 1
 DB: 1 Gaps: 1

US-09-940-235-10 (1-1661) x US-08-488-940-1 (1-1194)

QY	183	ATTGCTGGACCTGAGTGGCTGTAGACCTGCTCAACACAGCCCAATGGTGT	242
DB	382	IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal	401
QY	243	AGCGTTGTGCTGACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC	302
DB	402	SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle	421
QY	303	GATCTAACATCAGCCTGCTCATGAGGAAAGACAGACGAGCAAGCTTAAAGTCAAAATCA	362
DB	422	AspLeuThrSerArgProAlaHisGlyGlyLeuThrGluGlnGlyLeuSerProLysSer	441
QY	363	AAACCATTTGCTACTGATAGTGGCGGATGTACATAAACTGAGAAAGCTGACTACTA	422
DB	442	LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu	461
QY	423	AAGGCTATTCAAGAACAAATTGATCGCTAACGTCCACAGTAACGACGACTCTTTGAGGTC	482
DB	462	LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspIlePheGluVal	481
QY	483	ATTGATTTTGAAGCGATCAACCAATTACTGATCGAAACGGCAAGCTCTACTTTGCTGAC	542
DB	482	IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyPheAlaAsp	501
QY	543	AAAGATGTTTGGTAACTTGGCGGACCCCACTGTCCAGAAATTTTGTAAAGCGGACAT	602
DB	502	LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis	521
QY	603	GTGCGCGTTAGACCATATAAAGAAAACCAATACAAACCAAGCGAAATCTGTGATGTG	662
DB	522	ValArgValArg---TyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal	540
QY	663	GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAA	722
DB	541	GluTyThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys	560
QY	723	GATACCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCACTCTCAAGAATTACTA	782
DB	561	AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu	580

QY	783	GCTCAAGCACAAAGCATTTTAAACAAAAACCAACCCAGCTATACGATTATCAACGTCAC	842
DB	581	AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyThrIleTyGluArgAsp	600
QY	843	TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTTACCAATGATCAAGAG	902
DB	601	SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu	620
QY	903	TTTACTTACCGTGTATAAAATCGGAAACAAAGCTTATAGGATCAATAAAAAATCTGCTCTG	962
DB	621	PheThrTyArgValLysAsnArgGluGlnAlaTyArgIleAsnLysLysSerGlyLeu	640
QY	963	AATGAAGAAATAAAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAGGG	1022
DB	641	AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyTyTyValLeuLysLysGly	660
QY	1023	GAAGAAGCGTATGATCCCTTTGATCGGAGTCACCTTGAAACCTGTTCAACATCAATACGTT	1082
DB	661	GluLysProTyAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyVal	680
QY	1083	GATGTCGATACCAACGAATTGCTTAAAGTGACGAGCTCTTAAACAGCTAGCGAAGCTAAC	1142
DB	681	AspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn	700
QY	1143	TTAGACTTCAGAGATTTTATACGATCCCTGCTGATAGGCTAAAGCTACTCTACAACAATCTC	1202
DB	701	LeuAspPheArgAspLeuTyAspProArgAspLysAlaLysLeuLeuTyAsnAsnLeu	720
QY	1203	GATGCTTTTGGTATTATGACTATACCTTAACTGGAAGTAGAGGATAATCAGCATGAC	1262
DB	721	AspAlaPheGlyIleMetAspTyThrLeuThrGlyLysValGluAspAsnHisAspAsp	740
QY	1263	ACCAACCGTATCAACCGTTTATATGCGCAAGCGACCGAGAGAGAGATGCTAGCTAT	1322
DB	741	ThrAsnArgIleIleThrValTyMetGlyLysArgProGluGlyGluAsnAlaSerTy	760
QY	1323	CATTAGGCC 1331	
DB	761	HisLeuAla 763	

RESULT 8
 US-07-854-596B-26
 ; Sequence 26, Application US/07854596B
 ; Patent No. 5434073
 ; GENERAL INFORMATION:
 ; APPLICANT: Dawson, Keith M
 ; APPLICANT: Hunter, Michael G
 ; APPLICANT: Czaplowski, Lloyd G
 ; TITLE OF INVENTION: Proteins and nucleic acids
 ; NUMBER OF SEQUENCES: 73
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Dr. John J. McDonnell
 ; STREET: Ten South Wacker Drive, Suite 3000
 ; CITY: Chicago
 ; STATE: IL
 ; COUNTRY: USA
 ; ZIP: 60606
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/854,596B
 ; FILING DATE: 03-JUN-1992
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: McDonnell, John J
 ; REGISTRATION NUMBER: 26,949
 ; REFERENCE/DOCKET NUMBER: 92,337
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 312-715-1000
 ; TELEFAX: 312-715-1234

```

; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 415 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-854-596B-26

Alignment Scores:
Pred. No.: 1,02e-183 Length: 415
Score: 1963.00 Matches: 379
Percent Similarity: 98.96% Conservative: 1
Best Local Similarity: 98.70% Mismatches: 4
Query Match: 66.73% Indels: 0
DB: 1 Gaps: 0

US-09-940-235-10 (1-1661) x US-07-854-596B-26 (1-415)

QY 180 ATGATTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAAATGGTT 239
DB 1 MetIleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
QY 240 GTTAGCGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 299
DB 21 ValSerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLeuysPhePheGlu 40
QY 300 ATCGATCTAAATCATGACGACCTGCTCATGGAGGAGAGAGAGAGAGAGAGAGAGAGAG 359
DB 41 IleAspLeuThrSerArgProAlaHisGlyLysThrGluGlnGlyLeuSerProLys 60
QY 360 TCAAAACCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 419
DB 61 SerLysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGlnLysAlaAspLeu 80
QY 420 CTAAAGGCTATTCAAGAACAAATGATCGCTTAACGCTCCACAGTAAACGACGACTACTT 479
DB 81 LeuLysAlaIleGlnGlnLeuIleAlaValHisSerAsnAspAspThrPheGlu 100
QY 480 GTCAATGATTTGCAAGGATGCAACCATTAATGATGATGCAACGCGAGGCTGCTACTTGT 539
DB 101 ValIleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValThrPheAla 120
QY 540 GACAAAGATGCTCGGTACCTGCTCGGACCTGCTCGGACCTGCTCGGACCTGCTCGGAC 599
DB 121 AspLysAspGlySerValThrLeuProThrGlnProValGlnGlnPheLeuLeuSerGly 140
QY 600 CATGTGCGCGTTAGACCATATATAAGAAACCAATACAAACCAAGCGAAATCTGTTGAT 659
DB 141 HisValArgValArgProThrLysGluLysProIleGlnAsnGlnAlaLysSerValAsp 160
QY 660 GTGAATATATCTGACGTTTACTCCCTTAAACCCCTGATGACGATTCAGACGAGTCTC 719
DB 161 ValGluThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeu 180
QY 720 AAAGATATAGCTATTGAAACACTAGCTATCGGTGACACCATCATCATCTCAAGATTA 779
DB 181 LysAspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeu 200
QY 780 CTAGCTCAAGCACAAAGCATTTTAAACAAACCAACCAAGGCTATACGATTTTATGAACGT 839
DB 201 LeuAlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyThrThrIleThrGluArg 220
QY 840 GACTCCTCAATCGTCACTCATGACAAATGACATTTTCGGTACGATTTTACCAATGATCAA 899
DB 221 AspSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGln 240
QY 900 GAGTTTACTACCGTGTAAATAATCGGNAACAAGCTTATAGGATCATATAAAATCTGGT 959
DB 241 GluPheThrThrHisValLysAsnArgGluGlnAlaThrGluIleAsnLysLysSerGly 260
QY 960 CTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAA 1019

```

```

DB 261 LeuAsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrtYrValLeuLysLys 280
QY 1020 GGGGAAAAGCGGATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCCACCATCAATAC 1079
DB 281 GlyGluLysProTyrtYrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTy 300
QY 1080 GTTGATGTCGATACCAACGAATTTGCTAAAAAGTCAGCAGCTCTTAAACAGCTAGCGAAC 1139
DB 301 ValAspValAsnThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArg 320
QY 1140 AACTTAGCTTCAGAGATTTATACGATCTCTCGTGATAAGCTAACTACTCTACAAACAT 1199
DB 321 AsnLeuAspPheArgAspLeuTyrtYrAspProArgAspLysAlaLysLeuLeuTyrtYrAsn 340
QY 1200 CTCGATGCTTTTGGTATTATGAGTACTATACCTTAACTGGAAGAGTAGAGGATAATCAG 1259
DB 341 LeuAspAlaPheGlyIleMetAspTyrtYrThrLeuThrGlyLysValGluAspAsnHisAsp 360
QY 1260 GACACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAGGAGAGAGATGCTAGC 1319
DB 361 AspThrAsnArgIleIleThrValTyrtYrMetGlyLysArgProGluGlyGluAsnAlaSer 380
QY 1320 TATCATTTAGCC 1331
DB 381 TyrHisLeuAla 384

RESULT 9
US-07-854-596B-19
; Sequence 19, Application US/07854596B
; Patent No. 5434073
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith M
; APPLICANT: Hunter, Michael G
; APPLICANT: Czaplowski, Lloyd G
; TITLE OF INVENTION: Proteins and nucleic acids
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. John J. McDonnell
; STREET: Ten South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/854,596B
; FILING DATE: 03-JUN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,337
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 435 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-854-596B-19

Alignment Scores:
Pred. No.: 3.2e-183 Length: 435
Score: 1958.00 Matches: 378
Percent Similarity: 98.96% Conservative: 1
Best Local Similarity: 98.69% Mismatches: 4

```

Query Match:	66.62%	Indels:	0
DB:	1	Gaps:	0
US-09-940-235-10 (1-1661) x US-07-854-596B-19 (1-435)			
QY	183	ATTGCTGGAGCTGAGTGGCTGCTAGACCGTCCATCTGTCACACAGCCCAATTCGGTTGTT	242
DB	22	IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal	41
QY	243	AGCGTTGCTGGTACTGTTGAGGGGAGCAATCAAGACATTAGTCTTAATATTTTGAATC	302
DB	42	SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPheGluIle	61
QY	303	GATCTAACATCAGCACTGCTCATGTGAGGAGAAAGACAGCAAGGCTTAAGTCCAAATCA	362
DB	62	AspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer	81
QY	363	AAACCAATTCCTACTGATAGTGGCGGATGTCATAACTTGAGAAAGCTGACTTACTA	422
DB	82	LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu	101
QY	423	AGGCTATTCAAGCAATGATCGCTTAACGTCCACAGTAACGACGACTACTTTGAGTC	482
DB	102	LysAlaIleGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal	121
QY	483	ATTGATTTCGAAGCGATGCAACCATTTACTGTGAAACGGCAAGGTCTACTTTGCTGAC	542
DB	122	IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp	141
QY	543	AAAGATGTTGCGTAACTTCGCGACCCCACTGTCGAAGAAATTTTTCGTAAGCGGACAT	602
DB	142	LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis	161
QY	603	GTGCGGTAGACCATATAAGAAAAACCAATACAAACCAAGCGAAATCTGTGATGTG	662
DB	162	ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal	181
QY	663	GAATATACGTGACAGTTTACTCCCTTAAACCCGTGATGACCATTTTCAGCCAGGCTCAAA	722
DB	182	GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys	201
QY	723	GATCTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTA	782
DB	202	AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu	221
QY	783	GCTCAAGCAAGCAATTTTAAACAAAACCCAGCGGTATACGATTTTATGAACGTGAC	842
DB	222	AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp	241
QY	843	TCCTCAATCGTCACTCATGACATGACATTTTCGTACGATTTTACCAATGGATCAAGAG	902
DB	242	SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu	261
QY	903	TTTACTTACCGTGTAAAATCGGAAACAGCTTATAGGATCAATAAAAAATCTGGTCTG	962
DB	262	PheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsnLysLysSerGlyLeu	281
QY	963	AATGAAGAAATAACCAACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAGGG	1022
DB	282	AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly	301
QY	1023	GAAGAAGCCGTATGATCCCTTTGATCCGAGTCACTTGAAACTGTTCCACATCAATAGCTT	1082
DB	302	GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal	321
QY	1083	GATGTCGATACCAAGCAATTCCTTAAAGTGACGACCTCTTACAGCTAGCGACGTAA	1142
DB	322	AspValAsnThrAsnGlnLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn	341
QY	1143	TTAGACTTCAGAGATTATACGATCCTCGTGATAAGCTAAACTACTCTCAACAATCTC	1202
DB	342	LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu	361
QY	1203	GATGCTTTTGGTATTATGGACTATACCTTAACTGAAAAGTAGAGATAATCACCATGAC	1262
DB	362	AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp	381
QY	1263	ACCAACCGTATCATCAACCGTTTATATGGGCAAGGACCCGAGGAGAGAATCTGCTAGCTAT	1322
DB	382	ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr	401
QY	1323	CATTTAGCC 1331	
DB	402	HisLeuAla 404	
RESULT 10			
US-07-854-596B-15			
; Sequence 15, Application US/07854596B			
; Patent No. 5434073			
; GENERAL INFORMATION:			
; APPLICANT: Dawson, Keith M			
; APPLICANT: Hunter, Michael G			
; APPLICANT: Czaplewski, Lloyd G			
; TITLE OF INVENTION: Proteins and nucleic acids			
; NUMBER OF SEQUENCES: 73			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Dr. John J. McDonnell			
; STREET: Ten South Wacker Drive, Suite 3000			
; CITY: Chicago			
; STATE: IL			
; COUNTRY: USA			
; ZIP: 60606			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: Patencin Release #1.0, Version #1.25			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/07/854,596B			
; FILING DATE: 03-JUN-1992			
; CLASSIFICATION: 435			
; ATTORNEY/AGENT INFORMATION:			
; NAME: McDonnell, John J			
; REGISTRATION NUMBER: 26,949			
; REFERENCE/DOCKET NUMBER: 92,337			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: 312-715-1000			
; TELEFAX: 312-715-1234			
; INFORMATION FOR SEQ ID NO: 15:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 440 amino acids			
; TYPE: amino acid			
; TOPOLOGY: linear			
; MOLECULE TYPE: protein			
US-07-854-596B-15			
Alignment Scores:			
Pred. No.:	3,22e-183	Length:	440
Score:	1958.00	Matches:	378
Percent Similarity:	98.96%	Conservative:	1
Best Local Similarity:	98.69%	Mismatches:	4
Query Match:	66.62%	Indels:	0
DB:	1	Gaps:	0
US-09-940-235-10 (1-1661) x US-07-854-596B-15 (1-440)			
QY	183	ATTGCTGGAGCTGAGTGGCTGCTAGACCGTCCATCTCTCAACACAGCCCAATTCGGTTGTT	242
DB	27	IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal	46
QY	243	AGCGTTCTCGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAATATTTTGAATC	302
DB	47	SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPheGluIle	66
QY	303	GATCTAACATCAGCACTGCTCATGTGAGGAGAAAGACAGACGCAAGCGCTTAAGTCCAAATCA	362

```
Db 67 AspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 86
QY 363 AAACCATTTGCTACTGATAGTGGCGGATGTACATAAACTCAGAAAGCTGACTACTA 422
Db 87 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGlnLysAlaAspLeuLeu 106
QY 423 AAGGCTATTCAAGAACAAATTGATCGCTAAGCTCCACAGTAAACGACGACTACTTTGAGGTC 482
Db 107 LysAlaIleGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 126
QY 483 ATTGATTTTGCAGCGATCAACCAATTACTGATCGAAACGGCAAGGCTCTACTTTGCTGAC 542
Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146
QY 543 AAAGATGTTTGGTAACTTGGGACCCAACTGTCTCAAGAAATTTTGTCTAAGCGGACAT 602
Db 147 LysAspGlySerValThrLeuProThrGlnProValGlnGlnPheLeuLeuSerGlyHis 166
QY 603 GTGCGCGTTAGACCATATAAAGAAAAACCAATACAAAAACCAAGCGAAATCTGTTGATGTC 662
Db 167 ValArgValArgProTyrLysGlnLysProIleGlnAsnGlnAlaLysSerValAspVal 186
QY 663 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAAA 722
Db 187 GluTyrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuLys 206
QY 723 GATACTAGCTATTGAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTAATA 782
Db 207 AspThrLysLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeuLeu 226
QY 783 GCTCAAGCACAAAGCATTTTAAACAAAAACCCAGGCTATACGATTTATGACGCTGAC 842
Db 227 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp 246
QY 843 TCCTCAATCGTCACATGACATGACATTTTCGTTACGATTTTACCAATGGAATCAAGAG 902
Db 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
QY 903 TTTACTTACCGTGTAAATAACCGGAACAAAGCTTATAGGATCAATAAAAAATCTGCTCTG 962
Db 267 PheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsnLysLysSerGlyLeu 286
QY 963 AATGAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAAGTCTCTTAAAAAGGG 1022
Db 287 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 306
QY 1023 GAAAGCCGTATGATCCCTTTGATCGCTGCTGAACTGTTCAACCATCAAAATACGTT 1082
Db 307 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326
QY 1083 GATGTCGATACCAACGAATTGCTAAAGTAGCAGCTCTTAACAGCTAGCGCAACGTAAC 1142
Db 327 AspValAsnThrAsnGlnLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
QY 1143 TTAGACTTCAGAGATTATACATCTCTGTATAGGCTAAACTACTCTCAACAATCTC 1202
Db 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
QY 1203 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAAGTAGAGGATAATCACGATGAC 1262
Db 367 AspAlaPheGlyIleTetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 386
QY 1263 ACCAACCGTATCATACCGTTTATATGGCGAAGCGACCCGAAAGGAGAGAAATGCTAGCTAT 1322
Db 387 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 406
QY 1323 CATTTAGCC 1331
Db 407 HisLeuAla 409
```

RESULT 11

US-07-854-596B-43

```
: Sequence 43, Application US/07854596B
: Patent No. 5434073
: GENERAL INFORMATION:
: APPLICANT: Dawson, Keith M
: APPLICANT: Hunter, Michael G
: APPLICANT: Czaplowski, Lloyd G
: TITLE OF INVENTION: Proteins and nucleic acids
: NUMBER OF SEQUENCES: 73
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Dr. John J. McDonnell
: STREET: Ten South Wacker Drive, Suite 3000
: CITY: Chicago
: STATE: IL
: COUNTRY: USA
: ZIP: 60606
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/854,596B
: FILING DATE: 03-JUN-1992
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: McDonnell, John J
: REGISTRATION NUMBER: 26,949
: REFERENCE/DOCKET NUMBER: 92,337
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 312-715-1000
: TELEFAX: 312-715-1234
: TELEX: 910-221-5317
: INFORMATION FOR SEQ ID NO: 43:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 483 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-07-854-596B-43
```

```
Alignment Scores:
Pred. No.: 3 34e-183 Length: 483
Score: 1958.00 Matches: 378
Percent Similarity: 98.96% Conservative: 1
Best Local Similarity: 98.69% Mismatches: 4
Query Match: 66.62% Indels: 0
DB: 1 Gaps: 0
```

US-09-940-235-10 (1-1661) x US-07-854-596B-43 (1-483)

```
QY 183 ATGTGCGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAAATGGTTGTT 242
Db 70 IleAlaGlyProGluThrLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 89
QY 243 ACGGTTGCTGGTACTCTTGGGGGACGATCAACATAGACATTAAGTCTTAAATTTTGAATC 302
Db 90 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 109
QY 303 GATCTAACATCACGACCTGCTCATGGAGGAAACAGACAGCAAGCTTAAAGTCCAAAATCA 362
Db 110 AspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 129
QY 363 AAACCAATTTGCTACTGATAGTGGCGGATGTCAATAAACTTGAAGAGCTGACTTACTA 422
Db 130 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu 149
QY 423 AAGGCTATTCAAGAACAAATTGATCGCTAAGCTCCACAGTAAACGACGACTACTTTGAGGTC 482
Db 150 LysAlaIleGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 169
QY 483 ATTGATTTTGCAGCGATGCAACCAATTACTGATCGAAACCGCAAGGCTCTACTTTGCTGAC 542
Db 170 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 189
```

```

QY 543 AAAGATGTTTCGGTAACTTGGCGACCCCAACCTGTCCAAGAAATTTTGTCTAAGCGGACAT 602
Db 190 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 209
QY 603 GTGCGGTTAGACCATATAAGAAAAACCAATACAAACCAACCAAGCAATCTGTTGATGTG 662
Db 210 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 229
QY 663 GAATATACCTGTCAGTTTACTCCCTTAAACCCCTGATGACCATTTTCAGACCAGTCTCAAA 722
Db 230 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 249
QY 723 GATACCTAAGCTATTGAAAACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTACTA 782
Db 250 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 269
QY 783 GCTCAAGCAAAAGCMTTAAACAAACCAACCCAGGCTATACGATTTATGACATGAC 842
Db 270 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp 289
QY 843 TCCTCAATCGTCACTCATGACATGACATTTTCGGTACGATTTTACCAATGGATCAAGAG 902
Db 290 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 309
QY 903 TTTACTTACCGTGTAAATAATCGGAAACAGCTTATAGGATCAATAAAAAATCTGCTCTG 962
Db 310 PheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsnLysLysSerGlyLeu 329
QY 963 AATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1022
Db 330 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 349
QY 1023 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCCACCAATACGTT 1082
Db 350 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 369
QY 1083 GATGTCGATACCAACCAATGTCTAAAGTGAAGCAGCTCTTAAACAGTACGCAACGTAAC 1142
Db 370 AspValAsnThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 389
QY 1143 TTAGACTTCAGAGATTTATACATCTCTCGTATGATGATGATGATGATGATGATGATGAT 1202
Db 390 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 409
QY 1203 GATGCTTTTGGTATTATGACATATACCTTAACCTGGAAGAGTAGAGGATATACGATGAC 1262
Db 410 AspAlaPheGlyMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 429
QY 1263 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAGGAGAGATGCTAGCTAT 1322
Db 430 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 449
QY 1323 CATTTAGCC 1331
Db 450 HisLeuAla 452

```

```

RESULT 12
US-07-854-596B-47
; Sequence 47, Application US/07854596B
; Patent No. 5434073
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith M
; APPLICANT: Hunter, Michael G
; APPLICANT: Czaplowski, Lloyd G
; TITLE OF INVENTION: Proteins and nucleic acids
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. John J. McDonnell
; STREET: Ten South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA

```

```

; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/854,596B
; FILING DATE: 03-JUN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,337
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 483 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-854-596B-47

Alignment Scores:
Pred. No.: 3,34e-183 Length: 483
Score: 1958.00 Matches: 378
Percent Similarity: 98.96% Conservative: 1
Best Local Similarity: 98.69% Mismatches: 4
Query Match: 66.62% Indels: 0
DB: 1 Gaps: 0

US-09-940-235-10 (1-1661) x US-07-854-596B-47 (1-483)
QY 183 ATTGCTGGACCTCAGTGGCTGCTAGACCGTCCATCTCTCAACAACAGCAATTTGGTTGT 242
Db 1 IleAlaGlyProGluTyrPheLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
QY 243 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATATTAGTCTTAATATTTTGAATC 302
Db 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 40
QY 303 GATCTAACATCAGACCTGCTCATGAGGAGAAAGACAGACAGAGCTTAAGTCCAAAATCA 362
Db 41 AspLeuThrSerArgProAlaHisGlyLysThrGluGlnGlyLeuSerProLysSer 60
QY 363 AAACCATTTGCTACTGATAGTGGCGGATGTCACATAAACTTGAGAAAGCTGACTTACTA 422
Db 61 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu 80
QY 423 AAGCGTATTCAAGAACAAATTTGATCGCTTAACGTCACAGTAAACGACGACTACTTTGAGGTC 482
Db 81 LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100
QY 483 ATTGATTTTCAAGCGGATGCAACATTTACTGATCGAAACCGCAAGGCTCTACTTTGCTGAC 542
Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
QY 543 AAAGATGTTTCGGTAAACCTTGGCGACCCCAACCTGTCCAAGAAATTTTGTCTAAGCGGACAT 602
Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
QY 603 GTGCGGTTAGACCATATAAGAAAAACCAATACAAACCAACCAAGCAATCTGTTGATGTG 662
Db 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
QY 663 GAATATACCTGTCAGATTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCAAA 722
Db 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
QY 723 GATACCTAAGCTATTGAAAAACCACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTACTA 782

```



```

181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 200
783 GCTCAGCACAAGCATTTTAAACAAAACCCAGGCTATACGATTTATGACGTCAC 842
201 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp 220
843 TCCTCAATCGTCACATGACATGACATTTCCGTCACGATTTTACCAATGGATCAAGAG 902
221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
903 TTTACTTACCGTGTAAATAACCGGAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 962
241 PheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsnLysLysSerGlyLeu 260
963 AATGAAGAATAAACAACACTGACCTGATCTCTGAGAATATATACGTCCTTAAAAAGGG 1022
261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280
1023 GAAAGCCGATGATCCCTTTCATCGCAGTCACTTCAAACTGTTCACCATCAATACGTT 1082
281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
1083 GATGTCGATACCAACGAATGTCTAAAGCTGAGCAGCTCTTAACAGCTAGCGAACGTAAC 1142
301 AspValAsnThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320
1143 TTAGACTTCAGAGATTTATACATCTCTGTGATAGGCTAAACTACTCTACAAATCTC 1202
321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
1203 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAGAGTAGAGGATATACACGATGAC 1262
341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
1263 ACCAACCGTATCATACCGCTTTATATGGCAAGCGACCCGGAAGGAGAGAGATGCTAGCTAT 1322
361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 380
1323 CATTTAGCC 1331
381 HisLeuAla 383

```

RESULT 13

```

US-07-854-596B-28
; Sequence 28, Application US/07854596B
; Patent No. 5434073
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith M
; APPLICANT: Hunter, Michael G
; APPLICANT: Czaplowski, Lloyd G
; TITLE OF INVENTION: Proteins and nucleic acids
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. John J. McDonnell
; STREET: Ten South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/854,596B
; FILING DATE: 03-JUN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,337

```

TELECOMMUNICATION INFORMATION:

```

; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 499 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-854-596B-28

```

```

Alignment Scores:
Pred. No.: 3 38e-183 Length: 499
Score: 1958.00 Matches: 378
Percent Similarity: 98.96% Conservative: 1
Best Local Similarity: 98.69% Mismatches: 4
Query Match: 66.62% Indels: 0
DB: 1 Gaps: 0

```

US-09-940-235-10 (1-1661) x US-07-854-596B-28 (1-499)

```

QY 183 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCCAATGGTGT 242
Db 86 IleAlaGlyProGluThrLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 105
QY 243 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAATTTTGAATC 302
Db 106 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPheGluIle 125
QY 303 GATCTTAACATCACGACCTGCTCATGAGGAGAAACAGACAGCAAGCTTAAAGTCCAAAATCA 362
Db 126 AspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 145
QY 363 AAACCATTTGCTACTGATAGTGGCGCATGCTCATATAAACTTGAGAAAGCTGACTTACTA 422
Db 146 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu 165
QY 423 AAGGCTATTCAAGAACATTTGATCGCTTAACGTCACAGTAAACGACGACTACTTTGAGTGC 482
Db 166 LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 185
QY 483 ATTGATTTTCAAGCGATGCAACCATTTACTGATCGAAACCGCAAGGTCTACTTTGCTGAC 542
Db 186 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 205
QY 543 AAAGATGTTTCGGTAACTTTGCGACCCCAACCTGTCTCAAGAAATTTTGTAAAGCGGACAT 602
Db 206 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 225
QY 603 GTGCGGTTAGACCATATAAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 662
Db 226 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 245
QY 663 GAATATACGTACAGTTTACTCCCTTAAACCTGATGACCATTTTCAGACGAGCTCTCAA 722
Db 246 GluTyrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuLys 265
QY 723 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTA 782
Db 266 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 285
QY 783 GCTCAAGCACAAGCATTTTAAACAAAACCAAGCCAGGCTATACGATTTATGACGTCAC 842
Db 286 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp 305
QY 843 TCCTCAATCGTCACTCATGACATGACATTTCCGTCACGATTTTACCAATGGATCAAGAG 902
Db 306 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 325
QY 903 TTTTACTTACCGTGTAAATAACCGGAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 962
Db 326 PheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsnLysLysSerGlyLeu 345

```

QY 963 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 1022
Db 346 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrValLeuLysLysGly 365
QY 1023 GAAAGCGGTATGATCCCTTTGATCGGACGACCTGGAACCTGTCACCATCAATAGCTT 1082
Db 366 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 385
QY 1083 GATGTCGATACCAACGAATTTGCTAAAGAGTACGAGCTCTTAACAGCTAGCGACGTAAC 1142
Db 386 AspValAsnThrAsnGluLeuLysSerGluInLeuLeuThrAlaSerGluArgAsn 405
QY 1143 TTAGACTTCAGAGATTTATACGATCCCTCGTGAATAAGGCTAAACTACTCTACAAACAATCTC 1202
Db 406 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 425
QY 1203 GATGCTTTTGGTATTATGACTATACCTTAACCTGGAAAAAGTAGAGGATAATCAGCATGAC 1262
Db 426 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 445
QY 1263 ACCAACCGTATCATACCGTTTATATGGCGAAGCGCCGAGAGAGAAATGCTAGCTAT 1322
Db 446 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 465
QY 1323 CATTTAGGCC 1331
Db 466 HisLeuAla 468
RESULT 14
US-07-854-596B-35
; Sequence 35, Application US/07854596B
; Patent No. 5434073
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith M
; APPLICANT: Hunter, Michael G
; APPLICANT: Czaplowski, Lloyd G
; TITLE OF INVENTION: Proteins and nucleic acids
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Dr. John J. McDonnell
; STREET: Ten South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/854,596B
; FILING DATE: 03-JUN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,337
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 859 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-854-596B-35
Alignment Scores: 4.19e-183 Length: 859
Pred. No.:

Score: 1958.00 Matches: 378
Percent Similarity: 98.96% Conservative: 1
Best Local Similarity: 98.69% Mismatches: 4
Query Match: 66.62% Indels: 0
DB: 1 Gaps: 0
US-09-940-235-10 (1-1661) x US-07-854-596B-35 (1-859)
QY 183 ATTCTCGACCTGAGTGGCTCTAGACCGTCCATCTGTCAACAACAGCAATTTGGTCTGT 242
Db 22 IleAlaGlyProGluThrPleuLeuAspArgProSerValAsnAsnSerGlnLeuVal 41
QY 243 AGCGTTCGTCTACTGTTGAGGGACGAATCAAGACATTAGCTCTTAATATTTTGAATC 302
Db 42 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 61
QY 303 GATCTAATCATCAGACCTGCTCATGAGGAGAAAGACAGAGCAAGGCTTAAGTCCAAATCA 362
Db 62 AspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 81
QY 363 AAACCATTTCTACTAGTAGGCGCGATGTCACATAAATCTGAGAAGCTGACTTACTA 422
Db 82 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGluLysAlaAspLeu 101
QY 423 AAGCTATTCAAGAACCAATTCATCGCTAAACGTCACAGTAACGACGACTATCTTGAGTC 482
Db 102 LysAlaIleGlnGluInLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 121
QY 483 ATTGATTTTGAACGATGCAACCATCTACTGATCGAAACGCAAGGCTTACTTTGCTGAC 542
Db 122 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 141
QY 543 AAAGATGTTTCGGTAACTCTTCCCGACCCCACTGTCNAGAAATTTTCTAAGCGGACAT 602
Db 142 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 161
QY 603 GTGGCGCTTAGACCATATAAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATCTG 662
Db 162 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 181
QY 663 GAATATATCTGTACAGTTTACTTCCCTTAAACCTGATGACGATTTTCAGACCAAGGCTCAAA 722
Db 182 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 201
QY 723 GATCTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAATTACTA 782
Db 202 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 221
QY 783 GCTCAAGCACAAAGCATTTTAAACAAAACCCACCGCTATACGATTTATGAACGTGAC 842
Db 222 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp 241
QY 843 TCCTCAATCTGCTCACTCATGACAATGACATTTTCCGTACGATTTTACCAATCGATCAAGAG 902
Db 242 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 261
QY 903 TTACTTACCGTGTAAAAATTCGGAAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 962
Db 262 PheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsnLysSerGlyLeu 281
QY 963 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 1022
Db 282 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 301
QY 1023 GAAAGCGGTATGATCCCTTTGATCGCAGTCACTTTGAAACCTGTTCCACCATCAATAGCTT 1082
Db 302 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 321
QY 1083 GATGTCGATACCAACGAATTTGCTAAAAAGTAGAGAGCTCTTAAACAGCTAGCGACGTAAC 1142
Db 322 AspValAsnThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 341
QY 1143 TTAGACTTCAGAGATTTATAGGATCCCTCGTGAATAAGGCTAAACTACTCTACAAACAATCTC 1202

Db 342 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 361
Qy 1203 GATGCTTTTGGTATTATGACATACCTTAACCTGGAAAAGTAGAGGATAATCACGATGAC 1262
Db 362 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 381
Qy 1263 ACCAACCGTATCATACCGCTTTATATGGCAAGCGACCGAAGGAGAGATGCTAGCTAT 1322
Db 382 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 401
Qy 1323 CATTTAGCC 1331
Db 402 HisLeuAla 404

RESULT 15

US-08-488-940-17
; Sequence 17, Application US/08488940
; Patent No. 5854049
; GENERAL INFORMATION:
; APPLICANT: Reed, Guy L.
; TITLE OF INVENTION: PLASMIN-RESISTANT STREPTOKINASE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,940
; FILING DATE: 09-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 05433/009001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1194 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-488-940-17

Alignment Scores:
Pred. No.: 6,12e-181 Length: 1194
Score: 1936.50 Matches: 377
Percent Similarity: 98.43% Conservative: 0
Best Local Similarity: 98.43% Mismatches: 5
Query Match: 65.89% Indels: 1
DB: 1 Gaps: 1

US-09-940-235-10 (1-1661) x US-08-488-940-17 (1-1194)

Qy 183 ATTGCTGACCTGAGTGGTGTAGACCGCTCATCTGTCAACAGCCGAAATGGTGT 242
Db 382 IleAlaGlyProGluTrpLeuLeuAspAlaProSerValAsnAsnSerGlnLeuValVal 401
Qy 243 AGCGTTGCTGGTACCTGGAGGGGAGCAATCAAGACATTAGCTTAAATTTTGAATC 302
Db 402 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuAlaPhePheGluIle 421

Qy 303 GATCTAACATCATCGACCTGCTCATGGAGGAAAGACAGCAGCAAGCTTAAAGTCCAAATCA 362
Db 422 AspLeuThrSerAlaProAlaHisGlyGlyAlaThrGluGlnGlyLeuSerProAlaSer 441
Qy 363 AAACCATTTTCTACTAGTATGGCGCATGTACATAAAGCTTGAGAAAGCTGACTTACTA 422
Db 442 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 461
Qy 423 AAGCTATTCAAGAACAAATTGATCGCTAACGTCCACAGTAAACGACGACTACTTTGAGGTC 482
Db 462 LysAlaIleGlnGluLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 481
Qy 483 ATTGATTTTCCAAGCGATGCAACCATTTACTGATCGAAACCGCAAGGCTTACTTTGCTGAC 542
Db 482 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 501
Qy 543 AAAGATGGTTTCGGTAAACCTTTCGCCGACCAACCTGTCCAGAAATTTTTCGTAAGCGGACAT 602
Db 502 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 521
Qy 603 GTGCGGTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 662
Db 522 ValArgValArg--TyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 540
Qy 663 GAATATACTGTACAGTTTACTCCCTTAAACCCGTGATGACGATTTCAGACGAGTCTCAA 722
Db 541 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 560
Qy 723 GATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTA 782
Db 561 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 580
Qy 783 GCTCAAGCACAAGACATTTTAAACAAAACCAACCCAGGCTATACGATTTATGACGTCAC 842
Db 581 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 600
Qy 843 TCCTCAATCGTCACTCATGACAATGACATTTTCGTACGATTTTACCAATGGATCAAGAG 902
Db 601 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 620
Qy 903 TTTTACTTACCGTGTAAATTCGGACACGCTTATAGGATCAATAAAATCTGCTCTG 962
Db 621 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 640
Qy 963 AATGAAGAAATAAACCAACACTGACCTGATCTCTGAGAAATATTACGCTCTTAAAAAAGGG 1022
Db 641 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 660
Qy 1023 GAAAACCGGTATGATCCCTTTGATCGAGTCATCTTGAACCTGTTTCAACATCAATAGCTT 1082
Db 661 GluLysProTyrAspProPheAspArgSerHisLysLysLeuPheThrIleLysTyrVal 680
Qy 1083 GATGTCGATACCAACGAATTCCTTAAAGGTGACGACGCTCTTAACAGCTAGGACGTAAC 1142
Db 681 AspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 700
Qy 1143 TTAGACTTCAGAGATTATACGATCTCTCGTGATAAGGCTAAACTACTCTACAACTCTC 1202
Db 701 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 720
Qy 1203 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAAAGTAGAGGATAATACGATGAC 1262
Db 721 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 740
Qy 1263 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCGCCGAGAGAGATGCTAGCTAT 1322
Db 741 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 760
Qy 1323 CATTTAGCC 1331
Db 761 HisLeuAla 763

Search completed: January 28, 2006, 02:41:17
Job time : 42.6951 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 28, 2006, 01:51:16 ; Search time 108.622 Seconds
(without alignments)
12778.538 Million cell updates/sec

Title: US-09-940-235-10
Perfect score: 2939
Sequence: 1 gcaacccgcagcctagcc.....gaataagctgtaccatctaa 1661

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 3735138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US0940235/runat_27012006_144219_27635/app_query.fasta_1.7708
-DB=Published Applications AA Main -QFMT=fastan -SUFFIX=rapbm -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptco -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=200000000 -USER=US0940235_cgn_1_1_805_@runat_27012006_144219_27635
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DRV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-XGAPOP=6 -XGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA.Main.*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1984	67.5	413	4	US-10-360-101-264
2	1984	67.5	414	3	US-09-940-235-2
3	1963	66.8	415	5	US-10-988-943-1
4	1958	66.6	414	4	US-10-300-215-252
5	1947	66.2	414	4	US-10-300-215-253
6	1928	65.6	413	3	US-09-919-703-12
7	1794	61.0	440	5	US-10-474-792-658
8	600	20.4	259	3	US-09-940-235-4
9	600	20.4	2328	4	US-10-171-311-64
10	600	20.4	2328	4	US-10-236-031B-70
11	600	20.4	2328	4	US-10-374-979-98
12	600	20.4	2328	4	US-10-182-936A-98

13	600	20.4	2328	5	US-10-477-238A-677	Sequence 677, App
14	600	20.4	2328	5	US-10-680-287A-677	Sequence 677, App
15	600	20.4	2328	5	US-10-477-173-677	Sequence 677, App
16	600	20.4	2474	5	US-10-450-763-52638	Sequence 52638, A
17	596	20.3	231	3	US-09-925-302-548	Sequence 548, App
18	596	20.3	231	3	US-09-925-302-548	Sequence 548, App
19	596	20.3	463	4	US-10-144-194A-52	Sequence 52, Appl
20	596	20.3	463	5	US-10-491-566-52	Sequence 52, Appl
21	596	20.3	642	4	US-10-741-601-354	Sequence 354, App
22	596	20.3	642	5	US-10-741-600-1066	Sequence 1066, Ap
23	596	20.3	657	4	US-10-741-601-359	Sequence 359, App
24	596	20.3	657	5	US-10-741-600-1072	Sequence 1072, Ap
25	596	20.3	984	4	US-10-741-601-356	Sequence 356, App
26	596	20.3	984	5	US-10-741-600-1069	Sequence 1069, Ap
27	596	20.3	1173	5	US-10-450-763-52634	Sequence 52634, A
28	596	20.3	2220	4	US-10-236-392-4	Sequence 4, Appli
29	596	20.3	2296	4	US-10-741-601-363	Sequence 363, App
30	596	20.3	2296	5	US-10-741-600-1075	Sequence 1075, Ap
31	596	20.3	2355	4	US-10-144-194A-104	Sequence 104, App
32	596	20.3	2355	4	US-10-360-101-235	Sequence 235, App
33	596	20.3	2355	4	US-10-447-161-3	Sequence 3, Appli
34	596	20.3	2355	4	US-10-734-564-94	Sequence 94, Appl
35	596	20.3	2355	4	US-10-741-601-357	Sequence 357, App
36	596	20.3	2355	4	US-10-741-601-366	Sequence 366, App
37	596	20.3	2355	5	US-10-491-566-104	Sequence 104, App
38	596	20.3	2355	5	US-10-741-600-1067	Sequence 1067, Ap
39	596	20.3	2355	5	US-10-741-600-1078	Sequence 1078, Ap
40	596	20.3	2355	5	US-10-852-335A-147	Sequence 147, App
41	596	20.3	2355	5	US-10-287-436A-436	Sequence 436, App
42	596	20.3	2355	5	US-10-287-436A-1137	Sequence 1137, Ap
43	596	20.3	2355	6	US-11-040-130-28	Sequence 28, Appli
44	596	20.3	2386	3	US-09-961-403-1	Sequence 1, Appli
45	596	20.3	2386	4	US-10-741-601-360	Sequence 360, App

ALIGNMENTS

RESULT 1
US-10-360-101-264
; Sequence 264, Application US/10360101
; Publication No. US20040009550A1
; GENERAL INFORMATION:
; APPLICANT: Moll, Gert N.
; APPLICANT: Leenhouts, Cornelis J.
; TITLE OF INVENTION: Export and modification of (poly)peptide in the lantibiotic way
; FILE REFERENCE: 2183-5673
; CURRENT APPLICATION NUMBER: US/10/360,101
; CURRENT FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: EP 02077060.8
; PRIOR FILING DATE: 2002-05-24
; NUMBER OF SEQ ID NOS: 309
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 264
; LENGTH: 413
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: sequence of streptokinase
US-10-360-101-264

Alignment Scores:
Pred. No.: 1.06e-179 Length: 413
Score: 1984.00 Matches: 383
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 67.51% Indels: 0
DB: 4 Gaps: 0

US-09-940-235-10 (1-1661) x US-10-360-101-264 (1-413)

Qy 183 ATTGCTGGACCTCAGTGGCTGCTAGACCGCTCCATCTCTCAACACCCCACTGGTGTGT 242
Db 1 ILEAlGlyProGluIrpLeuLeuAepArgProSerValAanAenSerGlnLeuValVal 20

QY 243 AGCGTTGCTGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC 302
 Db |||||||
 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 40
 QY 303 GATCTAACATCAGCCTCTCATGAGGAGAAAGACAGACAGCGCTTAAGTCCAAATCA 362
 Db |||||||
 41 AspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 60
 QY 363 AAACCATTTGCTACTGATAGTCGCGGATGTCACATAAACTTCAGAAAGCTGACTACTA 422
 Db |||||||
 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 80
 QY 423 AAGGCTATTCAAGAACAAATGATCGCTAAACGCTCCACAGTAACACGACTACTTTGAGTTC 482
 Db |||||||
 81 LysAlaIleGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100
 QY 483 ATGATTTTGAAGGATCAACCACTACTGATCGAAACCGGACGAGGCTACTTTGCTGAC 542
 Db |||||||
 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
 QY 543 AAAGATGTTTCCGTAACCTTCGCGACCCCACTGTCACAGAAATTTTGTCTAAGCGGACAT 602
 Db |||||||
 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
 QY 603 GTGCGGTTTAGACCATATAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 662
 Db |||||||
 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
 QY 663 GAATATACTGTACGTTTACTCCCTTAAACCCCTGATGACGATTTACAGACGAGTCTCAA 722
 Db |||||||
 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
 QY 723 GATCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTTACTTA 782
 Db |||||||
 181 AspThrLysLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 200
 QY 783 GCTCAAGCACAAAGCATTTTAAACAAAACCCAGGCTATACGATTTTATGAACTGAC 842
 Db |||||||
 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 220
 QY 843 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 902
 Db |||||||
 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
 QY 903 TTTACTTACCGTGTAAATCGGAAACAAAGCTTATAGGATCAATATAAATACTGCTGTG 962
 Db |||||||
 241 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 260
 QY 963 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACCTCTTAAAAAGGG 1022
 Db |||||||
 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280
 QY 1023 GAAAAGCGTATGATCCCTTTGATCGCAGTCACTTGAACACTGTTTCAACCATCAATACGTT 1082
 Db |||||||
 281 GluLysProTyrAspProPheAspArgSerHisLysLeuLysLeuPheThrIleLysTyrVal 300
 QY 1083 GATGTCGATACCAACGAATGTCTAAAGTAGAGAGCTCTTAAACAGCTACGCAACGTAAC 1142
 Db |||||||
 301 AspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320
 QY 1143 TTAGACTTCAGAGATTATACCATCTCTGATAGGCTAAGCTAACTACTCTACACATCTC 1202
 Db |||||||
 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
 QY 1203 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAGAGTAGAGGATATACCATGAC 1262
 Db |||||||
 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
 QY 1263 ACCAACCGTATCAATACCGTTTATATGGGCAAGCGACCCGAGGAGAGAAATGCTAGCTAT 1322
 Db |||||||
 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 380

QY 1323 CATTTAGCC 1331
 Db |||||||
 381 HisLeuAla 383
 RESULT 2
 US-09-940-235-2
 ; Sequence 2, Application US/09940235
 ; Publication No. US20030059921A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kumar, Rajesh
 ; APPLICANT: Sahni, Girish
 ; APPLICANT: Roy, Chait
 ; APPLICANT: Rajagopal, Kammara
 ; APPLICANT: Nihalani, Deepak
 ; APPLICANT: Sundaram, Vasudha
 ; APPLICANT: Yadav, Mahavir
 ; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
 ; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
 ; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
 ; FILE REFERENCE: 07064-009002
 ; CURRENT APPLICATION NUMBER: US/09/940,235
 ; CURRENT FILING DATE: 2002-04-09
 ; PRIOR APPLICATION NUMBER: 09/471,349
 ; PRIOR FILING DATE: 1999-12-23
 ; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
 ; PRIOR FILING DATE: 1998-12-24
 ; NUMBER OF SEQ ID NOS: 28
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 414
 ; TYPE: PRT
 ; ORGANISM: Streptococcus equisimilis
 US-09-940-235-2
 Alignment Scores:
 Pred. No.: 1,06e-179 Length: 414
 Score: 1984.00 Matches: 383
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 67.51% Indels: 0
 DB: 3 Gaps: 0
 US-09-940-235-10 (1-1661) x US-09-940-235-2 (1-414)
 QY 183 ATTGCTGGACCTGAGTGGCTGTAGACCGCTCATCTGTCAACACAGCCAAATGTTGTT 242
 Db |||||||
 1 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
 QY 243 AGCGTTGCTGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC 302
 Db |||||||
 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 40
 QY 303 GATCTAACATCAGCCTCTCATGAGGAGAAAGACAGACAGCGCTTAAGTCCAAATCA 362
 Db |||||||
 41 AspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 60
 QY 363 AAACCATTTGCTACTGATAGTCGCGGATGTCACATAAACTTCAGAAAGCTGACTACTA 422
 Db |||||||
 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 80
 QY 423 AAGGCTATTCAAGAACAAATGATCGCTAAACGCTCCACAGTAACACGACTACTTTGAGTTC 482
 Db |||||||
 81 LysAlaIleGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100
 QY 483 ATGATTTTGAAGGATCAACCACTACTGATCGAAACCGGACGAGGCTACTTTGCTGAC 542
 Db |||||||
 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
 QY 543 AAAGATGTTTCCGTAACCTTCGCGACCCCACTGTCACAGAAATTTTGTCTAAGCGGACAT 602
 Db |||||||
 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140

```

QY 603 GTGCGGTAGACCATATAAAGAAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 662
Db 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
QY 663 GAATATACCTGTCAGCTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGTCTCAA 722
Db 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
QY 723 GATACCTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTACTA 782
Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 200
QY 783 GCTCAAGCAACAAGCAATTTAAACAAAACCAACCCAGGCTATACGATTTATGAACGTGAC 842
Db 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 220
QY 843 TCCTCAATCGTCACTCATGACATGACATTTCCGTACGATTTTACCAGTGGATCAAGAG 902
Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
QY 903 TTTACTTACCGTGTAAAAATCGGAAACAAAGCTTATAGGATCAATAAAAAATCTGGTCTG 962
Db 241 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 260
QY 963 AATGAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 1022
Db 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280
QY 1023 GAAGAGCCGTATGATCCCTTTGATCGCAGTCACCTGAACTGTTCCACCATCAATACGTT 1082
Db 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
QY 1083 GATGTCGATACCAACCAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 1142
Db 301 AspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320
QY 1143 TTAGACTTCAGAGATTTATACGATCTCTCGTGAAGGCTAAAGCTACTCTCAACAATCTC 1202
Db 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
QY 1203 GATGCTTTGGTATTATGACTATACCTTAACTGGAAAAAGTAGAGGATAATCAAGATGAC 1262
Db 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
QY 1263 ACCAACCGTATCATACCGTTTATATGGCAAGGACCCGCAAGGAGAGATGCTAGCTAT 1322
Db 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGlyGluAsnAlaSerTyr 380
QY 1323 CATTTAGCC 1331
Db 381 HisLeuAla 383

```

RESULT 3

```

US-10-988-943-1
; Sequence 1, Application US/10988943
; Publication No. US20050176085A1
; GENERAL INFORMATION:
; APPLICANT: Center for Genetic Engineering and Biotechnology
; TITLE OF INVENTION: METHOD OF SELECTIVE PEPTIDE ISOLATION FOR THE
; TITLE OF INVENTION: IDENTIFICATION AND QUANTITATIVE ANALYSIS OF PROTEINS IN
; TITLE OF INVENTION: COMPLEX MIXTURES.
; FILE REFERENCE: Proteomics CU2003-269
; CURRENT APPLICATION NUMBER: US/10/988,943
; CURRENT FILING DATE: 2004-11-15
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 415
; TYPE: PRT
; ORGANISM: Streptococcus equisimilis
US-10-988-943-1
Alignment Scores:

```

```

Pred. No.: 1.06e-177 Length: 415
Score: 1963.00 Matches: 379
Percent Similarity: 98.96% Conservative: 1
Best Local Similarity: 98.70% Mismatches: 4
Query Match: 66.79% Indels: 0
Db: 5 Gaps: 0

US-09-940-235-10 (1-1661) x US-10-988-943-1 (1-415)

QY 180 ATGATTGTCGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACCAAGCGCAATTCGTT 239
Db 1 MetIleAlaGlyProGluTyrLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
QY 240 GTTAGCGTGTCTGCTACTGTTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTTCGAA 299
Db 21 ValSerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGlu 40
QY 300 ATCGATCTAACATCAACGACCTGCTCATGGAGAAAGACAGACGAGCTTAAGTCCAAAA 359
Db 41 IleAspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLys 60
QY 360 TCRAAACCATTTGCTACTGATAGTGGCGCGATGTCACATAAATTTGAGAAAGCTGACTTA 419
Db 61 SerLysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGluLysAlaAspLeu 80
QY 420 CTAAGAGCTATTCAAGAACCAATTCATCGCTAACGTCCACAGTAAACGACGACTACTTTTCG 479
Db 81 LeuLysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGlu 100
QY 480 GTCAATTGATTTGCAAGCGATGCAACCAATTAATCTGATCGAAACGGCAAGCTCTACTTTGCT 539
Db 101 ValIleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAla 120
QY 540 GACAAAGATGTTCCGTAAACCTTCGCGACCCCAACCTGTCGCAAGAAATTTTTCGTAAGCGGA 599
Db 121 AspLysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGly 140
QY 600 CATGTGCGCTTAGACCATATAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGAT 659
Db 141 HisValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAsp 160
QY 660 GTGAAATATACTGTACACAGTTTACTCCCTTAAACCTGATGACGATTTTCAGACAGGCTCTC 719
Db 161 ValGluTyrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeu 180
QY 720 AAGATATCTAAGCTATTGAAACCACTAGCTATCGGTGACACCATCATCTCAAGAATTA 779
Db 181 LysAspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 200
QY 780 CTAGCTCAAGCACAAAGCATTTTAAACAAAACCAACCCAGGCTATACGATTTATGAACGT 839
Db 201 LeuAlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArg 220
QY 840 GACTCTCTCAATCGTCACTCATGACATGACATTTTCGTCAGATTTTACCAATCGATCAA 899
Db 221 AspSerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGln 240
QY 900 GAGTTTACTTACCGTGTAAAAATTCGGGAAACCAAGCTTATAGGATCAATAAAAAATCTGGT 959
Db 241 GluPheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsnLysLysSerGly 260
QY 960 CTGATGAAGAAATAAACCAACACACGTGACCTGATCTCTGAGAAATATTACGCTTCTTAAAA 1019
Db 261 LeuAsnGluGluIleAsnAsnThrAspLeuLysSerGluLysTyrTyrValLeuLysLys 280
QY 1020 GGGGAAAGCCGTATGATCCCTTTGATCGGAGTCACCTTGAACCTGTTCAACCATCAATAC 1079
Db 281 GlyGluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyr 300
QY 1080 GTTGATGTCGATACCAACGAATTCCTAAAAAGTACGAGCTCTTAAACAGCTAGCGAAACGT 1139
Db 301 ValAspValAsnThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArg 320

```

```
Qy 1140 AACTTAGACTTCAGAGATTATACGATCCTCGTGAAGCTAAAGCTAACTACTCTACACAAT 1199
Db 321 AsnLeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsn 340
Qy 1200 CTCGATGCTTTTGGTATTATGACTATACCTTAACCTGGAAAGTAGAGGATAATCAGCAT 1259
Db 341 LeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAsp 360
Qy 1260 GACACCAACCGTATCATCAACCGTTTATATGGGCAAGCGACCCGAAGGAGAGATGCTAGC 1319
Db 361 AspThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSer 380
Qy 1320 TATCATTTAGCC 1331
Db 381 TyrHisLeuAla 384

RESULT 4
US-10-300-215-252
; Sequence 252, Application US/10300215
; Publication No. US20030153043A1
; GENERAL INFORMATION:
; APPLICANT: CARR, Francis Joseph
; APPLICANT: ADAIR, Fiona Suzanne
; APPLICANT: HAMILTON, Anita Anne
; APPLICANT: CARTER, Graham
; TITLE OF INVENTION: METHOD FOR THE PRODUCTION OF
; FILE REFERENCE: MER-104-Con.1
; CURRENT APPLICATION NUMBER: US/10/300,215
; CURRENT FILING DATE: 2002-11-20
; PRIOR APPLICATION NUMBER: US 09/438,136
; PRIOR FILING DATE: 1999-11-10
; PRIOR APPLICATION NUMBER: WO PCT/GB98/01473
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: GB 9710480.6
; PRIOR FILING DATE: 1997-05-21
; PRIOR APPLICATION NUMBER: GB 9716197.0
; PRIOR FILING DATE: 1997-07-31
; PRIOR APPLICATION NUMBER: GB 9725270.4
; PRIOR FILING DATE: 1997-11-28
; PRIOR APPLICATION NUMBER: GB 9807751.4
; PRIOR FILING DATE: 1998-04-14
; PRIOR APPLICATION NUMBER: US 60/067,235
; PRIOR FILING DATE: 1997-12-02
; NUMBER OF SEQ ID NOS: 254
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 252
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Streptococcus equisimilis
US-10-300-215-252

Alignment Scores:
Pred. No.: 3,18e-177 Length: 414
Score: 1958.00 Matches: 378
Percent Similarity: 98.96% Conservative: 1
Best Local Similarity: 98.69% Mismatches: 4
Query Match: 66.62% Indels: 0
DB: 4 Gaps: 0

US-09-940-235-10 (1-1661) x US-10-300-215-252 (1-414)

Qy 183 ATTGCTGACCTGAGTGGCTGTAGACCTGTCATCTGTCACCAACAGCCAAATTTGGTGT 242
Db 1 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal 20
Qy 243 AGCGTTGCTGCTACTGTTGAGGGGAGCAATCAAGACATTAGTCTTAATTTTGGAAATC 302
Db 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 40
Qy 303 GATCTAACATCACGACCTGCTCATGGAGGAAAGACAGACGAGCGCTTAAGTCCAAAATCA 362
Db 41 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 60
```

RESULT 5

US-10-300-215-253
; Sequence 253, Application US/10300215
; Publication No. US20030153043A1

```
Qy 363 AAACCAATTTGCTACTGATAGTGGCGCATGTACATAAACTTGAGAAAGCTACTTACTA 422
Db 61 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu 80
Qy 423 AAGCTATTCAAGAACAAATTGATCGCTAAACGCTCCACAGTAAACGACGACTACTTTGAGT 482
Db 81 LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100
Qy 483 ATTGATTTTGAAGCGATGCAACCATTTACTGATCGAAACGCAAGGCTACTTTGCTGAC 542
Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
Qy 543 AAAGATGTTTCGGTAACTCTGCCGACCAACCTGTCCAAGAAATTTTCTTAAGCGGCAT 602
Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
Qy 603 GTGCGCTTTAGACCATATAAAGAAAAACCAATACAAAAACCAAGCGAAATCTGTTGATGT 662
Db 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
Qy 663 GAATATACTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTCAGACCGAGTCTCAA 722
Db 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
Qy 723 GATACTAAGCTATTGAAAAACACTAGCTATCGTGCACACCATCATCATCTCAAGAAATTA 782
Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 200
Qy 783 GCTCAAGCACAAAGCATTTTAAACAAAAACCAACCCAGGCTATACGATTTATGAAACGT 842
Db 201 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp 220
Qy 843 TCCTCAATCGTCACTCATGACAATGACATTTTCGCTAGGATTTTACCAATGATCAAGAG 902
Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
Qy 903 TTTTACTTACCGTGTAAATAATCGGAAACACAGCTTATAGGATCAATAAAATCTGGTCTG 962
Db 241 PheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsnLysLysSerGlyLeu 260
Qy 963 AATGAAGAAATAAACACACACTGACCTGATCTCTGAGAAATATTAGCTCTTAAAAAGGG 1022
Db 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280
Qy 1023 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACCTTGAAACCTGTTCCACCATCAATAC 1082
Db 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
Qy 1083 GATGTCGATACCAACGAATTGCTTAAAGAGCAGCAGCTCTTAAACAGCTAGCGAAGTAA 1142
Db 301 AspValAsnThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320
Qy 1143 TTAGACTTCAGAGATTTTATACGATCCTCGTGATAGGCTAAAGCTACTCTACAAACATCTC 1202
Db 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
Qy 1203 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAGAGTAGAGGATAATCAGCATGAC 1262
Db 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
Qy 1263 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAAATGCTAGCTAT 1322
Db 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 380
Qy 1323 CATTTAGCC 1331
Db 381 HisLeuAla 383
```


GENERAL INFORMATION:
 APPLICANT: CARR, Francis Joseph
 APPLICANT: ADAIR, Fiona Suzanne
 APPLICANT: HAMILTON, Anita Anne
 APPLICANT: CARTER, Graham
 TITLE OF INVENTION: METHOD FOR THE PRODUCTION OF
 TITLE OF INVENTION: NON-IMMUNOGENIC PROTEINS
 FILE REFERENCE: MER-104-Cor.1
 CURRENT APPLICATION NUMBER: US/10/300,215
 CURRENT FILING DATE: 2002-11-20
 PRIOR APPLICATION NUMBER: US 09/438,136
 PRIOR FILING DATE: 1999-11-10
 PRIOR APPLICATION NUMBER: WO PCT/GB98/01473
 PRIOR FILING DATE: 1998-05-21
 PRIOR APPLICATION NUMBER: GB 9710480.6
 PRIOR FILING DATE: 1997-05-21
 PRIOR APPLICATION NUMBER: GB 9716197.0
 PRIOR FILING DATE: 1997-07-31
 PRIOR APPLICATION NUMBER: GB 9725270.4
 PRIOR FILING DATE: 1997-11-28
 PRIOR APPLICATION NUMBER: GB 9807751.4
 PRIOR FILING DATE: 1998-04-14
 PRIOR APPLICATION NUMBER: US 60/067,235
 PRIOR FILING DATE: 1997-12-02
 NUMBER OF SEQ ID NOS: 254
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 253
 LENGTH: 414
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Modified strep protein

US-10-300-215-253

Alignment Scores:
 Pred. No.: 3,55e-176 Length: 414
 Score: 1947.00 Matches: 376
 Percent Similarity: 98.69% Conservative: 2
 Best Local Similarity: 98.17% Mismatches: 5
 Query Match: 66.25% Indels: 0
 DB: 4 Gaps: 0

US-09-940-235-10 (1-1661) x US-10-300-215-253 (1-414)

QY 183 ATTGCTGGACCTGAGTGGCTGTAGACCTCCATCTGTCACACAGCCAAATGGTGT 242
 Db 1 IleAlaGlyProGluTrpLeuLeuAsePrgProSerValAsnAsnSerGlnLeuVal 20
 QY 243 AGCGTTGCTGCTGCTGCTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC 302
 Db 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspPheLeuLysPheGluIle 40
 QY 303 GATCTAACATCAGACCTGCTCATGGAGGAAAGACAGCAAGGCTTAAGTCCAAATCA 362
 Db 41 AseLeuThrSerArgProAlaHisGlyGlyGlyThrGluGlnGlyLeuSerProLysSer 60
 QY 363 AAACATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 422
 Db 61 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGluLysAlaAseLeuLeu 80
 QY 423 AAGGCTATTCAAGAACAAATGGATCGCTAACGTCACAGTAACGACGACTACTTTGAGTTC 482
 Db 81 LysAlaLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 100
 QY 483 ATTGATTTTGAAGCGATCAACCATTTACTGATCGAAACGGCAAGGCTTACTTTGCTGAC 542
 Db 101 IleAsePheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValThrPheAlaAse 120
 QY 543 AAAGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 602
 Db 121 LysAseGlySerValThrLeuProThrGlnProValGlnGlnPheLeuLeuSerGlyHis 140
 QY 603 GTGCCGCTTAGACCATATATAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 662

Db 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
 QY 663 GAATATATCTGTACAGTTTACTCCCTTAAACCCGTGATGACGATTTTACAGACGAGTCTCAA 722
 Db 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
 QY 723 GATACTAAGCTATTGAAACACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTA 782
 Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 200
 QY 783 GCTCAAGCACAAGCATTTTAAACAAACACCCAGGCTATACGATTTATGAACGTGAC 842
 Db 201 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp 220
 QY 843 TCCTCAATCGTCACTCATGACATGACATTTTCGTAGCATTTTACCAATGGATCAAGAG 902
 Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
 QY 903 TTTACTTACCGTGTAAAAATCGGAAACAAGCTTATAGGATCAATAAAAAATCTGTCGTG 962
 Db 241 PheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsnLysLysSerGlyLeu 260
 QY 963 AATCAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 1022
 Db 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280
 QY 1023 GAAAAGCGGTATGATCCCTTTGATCGCAGTCATCTGAAACTGTTCCACATCAATACGTT 1082
 Db 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysPheVal 300
 QY 1083 GATGTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAACAGCTACGACGTAAC 1142
 Db 301 AspValAsnThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320
 QY 1143 TTAGACTTCAGAGATTTATACGATCTCTCGTGAAGGCTAAAGCTAACTCTCAACAATCTC 1202
 Db 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
 QY 1203 GATGCTTTGCTGATTTATGACTATACCTTAACCTGGAAGAGTAGAGGATATACGATGAC 1262
 Db 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
 QY 1263 ACCAAGCGTATCATACCGTTTATATGGCAAGGACCCCAAGAGAGATGCTAGCTAT 1322
 Db 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 380
 QY 1323 CATTTAGGCC 1331
 Db 381 HisLeuAla 383

RESULT 6
 US-09-919-703-12
 Sequence 12, Application US/09919703
 Patent No. US20020165129A1
 GENERAL INFORMATION:
 APPLICANT: Krystal, Gerald
 APPLICANT: Rabkin, Simon W.
 TITLE OF INVENTION: Peptides and Their Use to Ameliorate
 TITLE OF INVENTION: Cell Death
 FILE REFERENCE: 50216/003004
 CURRENT APPLICATION NUMBER: US/09/919,703
 CURRENT FILING DATE: 2001-07-31
 PRIOR APPLICATION NUMBER: US 09/294,457
 PRIOR FILING DATE: 1999-04-19
 PRIOR APPLICATION NUMBER: US 08/759,599
 PRIOR FILING DATE: 1996-12-05
 PRIOR APPLICATION NUMBER: US 60/008,233
 PRIOR FILING DATE: 1995-12-06
 NUMBER OF SEQ ID NOS: 16
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 12
 LENGTH: 413

Db 280 GlyGluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyr 299

QY 1080 GTTGATGTGATACCAACGAATTCTTAAAAAGTCAGCAGCTCTTAAACAGCTAGCGAAGCT 1139

[illegible]

340 LEHAR, R. A. T. A. R. N. E. G. U. T. T. E. M. E. S. A. R. P. U. T. I. N. G. L. E. U. I. N. G. U. Y. U. S. V. A. I. G. T. A. N. S. P. R. A. S. I. N. I. S. A. R. P. 353

```

QY      1260 GATACCAACCGGATCATACCGGTTTATATAGCGCGACCGCGAGGAGAGATGCTAGC 131
Db      360 AspThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSer 379
QY      1320 TATCAT 1325
Db      380 TyrHis 381

RESULT 7
US-10-474-792-658
; Sequence 658, Application US/10474792
; Publication No. US20040236072A1
; GENERAL INFORMATION:
; APPLICANT: Olmsted, Stephen
; APPLICANT: Zagursky, Robert
; APPLICANT: Nickbarg, Elliot
; APPLICANT: Winter, Lourie
; TITLE OF INVENTION: SURFACE PROTEINS OF STREPTOCOCCUS PYOGENES
; FILE REFERENCE: AM 100399
; CURRENT APPLICATION NUMBER: US/10/474,792
; CURRENT FILING DATE: 2003-10-14
; NUMBER OF SEQ ID NOS: 674
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 658
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes
US-10-474-792-658

Alignment Scores:
Pred. No.: 1.36e-161 Length: 440
Score: 1794.00 Matches: 341
Percent Similarity: 93.99% Conservative: 19
Best Local Similarity: 89.03% Mismatches: 23
Query Match: 61.04% Indels: 0
DB: 5 Gaps: 0

US-09-940-235-10 (1-1661) x US-10-474-792-658 (1-440)
QY      183 ATTGCTGGACCTGAGTCGGCTGCTAGACCGTCCATCTCTCAACACAGCCCAATTGGTTGTT 242
Db      27 IleAlaGlyTyrGlyTrpLeuProAspArgProPheIleAsnAsnSerGlnLeuValVal 46
QY      243 AGCGTCTCGTACTGTTTGAGGGGACCAATTCACAGACATTAGTCTTAAATTTTTTGAATC 302
Db      47 SerMetAlaGlyIleValGluGlyThrAspLysLysValPheIleAsnPhePheGluIle 66
QY      303 GATCTAATCATCAGACCTGCTCATGGAGGAAAGACAGACAGCAGCGCTTAAGTCCAAATCA 362
Db      67 AspLeuThrSerGlnProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 86
QY      363 AAACCAATTGCTACTGATGTAGCGCGCATGCTCACATAAACTTCAGAAAGCTGACTTACTA 422
Db      87 LysProPheAlaThrAspAsnGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu 106
QY      423 AAGCCTATTCAAGAACAAATTGATCGCTTAACGCTCCACAGTACACGACGACTCTTTGAGGTC 482
Db      107 LysAlaIleGlnLysGlnLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal 126

```

```
QY 483 ATTGATTTTCGAAGCGATGCAACATTACTGATCGAAACGGCAAGGTCTACTTTGCTGAC 542
|||
Db 127 lIeAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146
|||
QY 543 AAAGATGTTTCGGTAACTTCGCCGACCAACCTGTCACGAATTTTTCGTAAGCGGCAT 602
|||
Db 147 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuLysGlyHis 166
|||
QY 603 GTGCGGTTAGACCATATAAGAAACCAACCAATACAAACCAACCAACCAATCTGTGATGTG 662
|||
Db 167 ValArgValArgProTyrLysGluLysProValGlnAsnGlnAlaLysSerValAspVal 186
|||
QY 663 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTACAGCCAGGTCCTCAA 722
|||
Db 187 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 206
|||
QY 723 GATACCTAAGCTATTGAACACTAGCTATCGGTGACACCATCATCTCAAGAACTACTA 782
|||
Db 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 226
|||
QY 783 GCTCAAGCAACAAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTATGACGTCAC 842
|||
Db 227 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp 246
|||
QY 843 TCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 902
|||
Db 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
|||
QY 903 TTTACTTACCGTGTGTTAAATTCGGGAACAGCTTTATAGATTCATATAAAATCTGGTCTG 962
|||
Db 267 PheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsnProLysThrGlyIle 286
|||
QY 963 AATGAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1022
|||
Db 287 LysGluLysThrAsnAsnThrAspLeuValSerGluLysTyrTyrValLeuLysGlnGly 306
|||
QY 1023 GAAAGCCGTATGATCCCTTTGATCGCAGTCACCTTGAAACTGTTTACCACCAATACGTT 1082
|||
Db 307 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326
|||
QY 1083 GATGCTGATCAACAAACGAATGCTTAAAGTGACAGCTCTTAAACAGTACGCAACGTAAC 1142
|||
Db 327 AspValAsnThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
|||
QY 1143 TTAGACTTCAGAGATTTATACGATCTCTGATGATAAGCTAACTACTCTCAACAATCTC 1202
|||
Db 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
|||
QY 1203 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAAAAGTAGAGGATATACGATGAC 1262
|||
Db 367 AspAlaPheAspIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspLys 386
|||
QY 1263 ACCAACCGTATCATACCGTTTATATGGCAACGACCGCAAGAGAGAGATGCTAGCTAT 1322
|||
Db 387 AsnAsnArgValValThrValTyrMetGlyLysArgProLysGlyAlaLysGlySerTyr 406
|||
QY 1323 CATTTAGCC 1331
|||
Db 407 HisLeuAla 409
|||
```

RESULT 8

```
US-09-940-235-4
; Sequence 4, Application US/09940235
; Publication No. US20030059921A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahn, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammara
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
```

```
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/09/940,235
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-940-235-4
Alignment Scores:
Pred. No.: 6,29e-48 Length: 259
Score: 600.00 Matches: 106
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.42% Indels: 0
DB: 3 Gaps: 0
US-09-940-235-10 (1-1661) x US-09-940-235-4 (1-259)
QY 1341 CAGCGCAGCAAAATGGTTTCAGCCCGGTCCTGTCAGTCAAAGCAAGCCCGT 1400
Db 1 GlnAlaGlnGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 20
QY 1401 TGTATGACAAATGAAAAACACTATCAGATAAATCAACGTGGGAGCGGACCTACCTAGGT 1460
Db 21 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTyrGluArgThrTyrLeuGly 40
QY 1461 AATGTGTTGGTTTGTACTTGTATTGGAGGAGCGAGGTTTAACTCGGAAAGTAAACCT 1520
Db 41 AsnValLeuValCysThrCysTyrGlySerArgGlyPheAsnGlnCysGlnSerLysPro 60
QY 1521 GAAGCTGAAGAGACTTCTTTTGACAAGTACACTGCGGAACACTTACCAGAGTGGGTGACACT 1580
Db 61 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
QY 1581 TATGAGCGTCTTAAAGACTCCATCATCTGGAGCTGTGACTGCTGATCGGGGCTGGGCGAGG 1640
Db 81 TyrGluArgProLysAspSerMetIleTyrAspCysThrCysIleGlyAlaGlyArgGly 100
QY 1641 AGATAAGCTGTACCATC 1658
Db 101 ArgIleSerCysThrIle 106
|||
RESULT 9
US-10-171-311-64
; Sequence 64, Application US/10171311
; Publication No. US20030087270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Glatt, Karen
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Hoersh, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
```

; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 64
; LENGTH: 2328
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-171-311-64

Alignment Scores:
Pred. No.: 1.3e-47 Length: 2328
Score: 600.00 Matches: 106
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.42% Indels: 0
DB: 4 Gaps: 0

US-09-940-235-10 (1-1661) x US-10-171-311-64 (1-2328)

QY 1341 CAGCGCAGCAAAATGTTTTCAGCCCGTCCCGTGGCTGTCTCAGTCAAGCAAGCCCGGT 1400
Db 5 GlnAlaGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 24
QY 1401 TGTATGACAATGGAACAACTATACAGATAAATCAACAGTGGAGCGGACCTACCTAGGT 1460
Db 25 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 44
QY 1461 AATGTGTGTTTGTCTTACTTGTATGAGGAGCGGAGGTTTAACTCGCAAGTAAACCT 1520
Db 45 AsnValLeuValCysThrCysTyrGlySerArgGlyPheAsnCysGluSerLysPro 64
QY 1521 GAAGCTGAAGAGACTTGTCTTTCAGCAAGTACACTGGGAAACACTTACCAGTGGGTGACACT 1580
Db 65 GluAlaGluGlnThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 84
QY 1581 TATGAGCGTCCTAAAGACTCCATGATCTGGGACTGTACCTGCATCGGGCTGGCGAGGG 1640
Db 85 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 104
QY 1641 AGAATAAGCTGTACCATC 1658
Db 105 ArgIleSerCysThrIle 110

RESULT 10

; Sequence 70, Application US/10236031B
; Publication No. US20030219760A1
; GENERAL INFORMATION:
; APPLICANT: Gordon, Gavin J.
; APPLICANT: Jensen, Roderick V.
; APPLICANT: Gullans, Steven R.
; APPLICANT: Bueno, Raphael
; TITLE OF INVENTION: Diagnostic and Prognostic Tests
; FILE REFERENCE: B00801/70265 (JRV/JAV)
; CURRENT APPLICATION NUMBER: US/10/236,031B
; CURRENT FILING DATE: 2002-09-05
; PRIOR APPLICATION NUMBER: US 60/317,389
; PRIOR FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: US 60/407,431
; PRIOR FILING DATE: 2002-08-30
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 70
; LENGTH: 2328
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-236-031B-70

Alignment Scores:
Pred. No.: 1.3e-47 Length: 2328
Score: 600.00 Matches: 106
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.42% Indels: 0
DB: 4 Gaps: 0

Pred. No.: 1.3e-47 Length: 2328
Score: 600.00 Matches: 106
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.42% Indels: 0
DB: 4 Gaps: 0

US-09-940-235-10 (1-1661) x US-10-236-031B-70 (1-2328)

QY 1341 CAGCGCAGCAAAATGTTTTCAGCCCGTCCCGTGGCTGTCTCAGTCAAGCAAGCCCGGT 1400
Db 5 GlnAlaGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 24
QY 1401 TGTATGACAATGGAACAACTATACAGATAAATCAACAGTGGAGCGGACCTACCTAGGT 1460
Db 25 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 44
QY 1461 AATGTGTGTTTGTCTTACTTGTATGAGGAGCGGAGGTTTAACTCGCAAGTAAACCT 1520
Db 45 AsnValLeuValCysThrCysTyrGlySerArgGlyPheAsnCysGluSerLysPro 64
QY 1521 GAAGCTGAAGAGACTTGTCTTTCAGCAAGTACACTGGGAAACACTTACCAGTGGGTGACACT 1580
Db 65 GluAlaGluGlnThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 84
QY 1581 TATGAGCGTCCTAAAGACTCCATGATCTGGGACTGTACCTGCATCGGGCTGGCGAGGG 1640
Db 85 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 104
QY 1641 AGAATAAGCTGTACCATC 1658
Db 105 ArgIleSerCysThrIle 110

RESULT 11

US-10-374-979-98
; Sequence 98, Application US/10374979
; Publication No. US20030219793A1
; GENERAL INFORMATION:
; APPLICANT: John P. Carulli et al.
; TITLE OF INVENTION: THE HIGH BONE MASS GENE OF 11q13.3
; FILE REFERENCE: 032796-021
; CURRENT APPLICATION NUMBER: US/10/374,979
; CURRENT FILING DATE: 2003-03-04
; PRIOR APPLICATION NUMBER: US 09/544,398
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 09/543,771
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 09/229,319
; PRIOR FILING DATE: 1999-01-13
; PRIOR APPLICATION NUMBER: US 60/071,449
; PRIOR FILING DATE: 1998-01-13
; PRIOR APPLICATION NUMBER: US 60/105,511
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 109
; SEQ ID NO 98
; LENGTH: 2328
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-374-979-98

Alignment Scores:
Pred. No.: 1.3e-47 Length: 2328
Score: 600.00 Matches: 106
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.42% Indels: 0
DB: 4 Gaps: 0

US-09-940-235-10 (1-1661) x US-10-374-979-98 (1-2328)

QY 1341 CAGCGCAGCAAAATGTTTTCAGCCCGTCCCGTGGCTGTCTCAGTCAAGCAAGCCCGGT 1400
Db 5 GlnAlaGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 24

```
Qy 1401 TGTATGACAAATGGAACACATATCATGATAAATCAACAGTGGGAGCGGACCTACCTAGGT 1460
Db 25 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 44
Qy 1461 AATGTGTGGTTGTACTTGTATGAGGAGCGGAGGTTTAACTGCGAAAGTAAACCT 1520
Db 45 AsnValLeuValCysTyrGlySerArgGlyPheAsnGlySerGlyPro 64
Qy 1521 GAAGCTGAAGAGACTTGTTCACAAAGTACACAGTGGGAACACTTACCGAGTGGGTGACACT 1580
Db 65 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 84
Qy 1581 TATGAGCGTCTTAAGACTCCATGATCTGGGACTGTACCTGCATCGGGCTGGGCGAGGG 1640
Db 85 TyrGluArgProLysAspSerMetIleTyrAspCysThrCysIleGlyAlaGlyArgGly 104
Qy 1641 AGAATAAGCTGTACCATC 1658
Db 105 ArgIleSerCysThrIle 110
RESULT 12
US-10-182-936A-98
; Sequence 98, Application US/10182936A
; Publication No.: US20040038860A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Kristina M.
; APPLICANT: Anisowicz, Anthony
; APPLICANT: Bhat, Bheem
; APPLICANT: Damagnez, Veronique
; APPLICANT: Robinson, John
; APPLICANT: Yaworsky, Paul
; TITLE OF INVENTION: Reagents and Method for Modulating DKK-Mediated Interactions
; CURRENT APPLICATION NUMBER: US/10/182,936A
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: PCT/US02/15982
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: US 60/291,311
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: US 60/353,058
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: US 60/361,293
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 216
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 98
; LENGTH: 2328
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-182-936A-98
Alignment Scores:
Pred. No.: 1,3e-47 Length: 2328
Score: 600.00 Matches: 106
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.42% Indels: 0
DB: 4 Gaps: 0
US-09-940-235-10 (1-1661) x US-10-182-936A-98 (1-2328)
Qy 1341 CAGGCGAGCAAAATGGTTCAGCCCGAGTCCCGGTGGCTGTCAAGTCAAAAGCAAGCCCGGT 1400
Db 5 GlnAlaGlnGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 24
Qy 1401 TGTATGACAAATGGAACACATATCATGATAAATCAACAGTGGGAGCGGACCTACCTAGGT 1460
Db 25 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 44
Qy 1461 AATGTGTGGTTGTACTTGTATGAGGAGCGGAGGTTTAACTGCGAAAGTAAACCT 1520
Db 45 AsnValLeuValCysTyrGlySerArgGlyPheAsnGlySerGlyPro 64
Qy 1521 GAAGCTGAAGAGACTTGTTCACAAAGTACACAGTGGGAACACTTACCGAGTGGGTGACACT 1580
Db 65 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 84
Qy 1581 TATGAGCGTCTTAAGACTCCATGATCTGGGACTGTACCTGCATCGGGCTGGGCGAGGG 1640
Db 85 TyrGluArgProLysAspSerMetIleTyrAspCysThrCysIleGlyAlaGlyArgGly 104
Qy 1641 AGAATAAGCTGTACCATC 1658
Db 105 ArgIleSerCysThrIle 110
```

```
Qy 1521 GAAGCTGAAGAGACTTGTTCACAAAGTACACAGTGGGAACACTTACCGAGTGGGTGACACT 1580
Db 65 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 84
Qy 1581 TATGAGCGTCTTAAGACTCCATGATCTGGGACTGTACCTGCATCGGGCTGGGCGAGGG 1640
Db 85 TyrGluArgProLysAspSerMetIleTyrAspCysThrCysIleGlyAlaGlyArgGly 104
Qy 1641 AGAATAAGCTGTACCATC 1658
Db 105 ArgIleSerCysThrIle 110
RESULT 13
US-10-477-238A-677
; Sequence 677, Application US/10477238A
; Publication No.: US20040221328A1
; GENERAL INFORMATION:
; APPLICANT: Babi, Philip
; APPLICANT: Yaworsky, Paul
; APPLICANT: Bex, Frederick J. III
; APPLICANT: Bodine, Peter Van Nest
; TITLE OF INVENTION: Transgenic Animal Model of Bone Mass Modulation
; FILE REFERENCE: 032796-212
; CURRENT APPLICATION NUMBER: US/10/477,238A
; CURRENT FILING DATE: 2003-11-10
; PRIOR APPLICATION NUMBER: US 60/290,071
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/291,311
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: US 60/353,058
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: US 60/361,293
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 812
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 677
; LENGTH: 2328
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-477-238A-677
Alignment Scores:
Pred. No.: 1,3e-47 Length: 2328
Score: 600.00 Matches: 106
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.42% Indels: 0
DB: 5 Gaps: 0
US-09-940-235-10 (1-1661) x US-10-477-238A-677 (1-2328)
Qy 1341 CAGGCGAGCAAAATGGTTCAGCCCGAGTCCCGGTGGCTGTCAAGTCAAAAGCAAGCCCGGT 1400
Db 5 GlnAlaGlnGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 24
Qy 1401 TGTATGACAAATGGAACACATATCATGATAAATCAACAGTGGGAGCGGACCTACCTAGGT 1460
Db 25 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 44
Qy 1461 AATGTGTGGTTGTACTTGTATGAGGAGCGGAGGTTTAACTGCGAAAGTAAACCT 1520
Db 45 AsnValLeuValCysTyrGlySerArgGlyPheAsnGlySerGlyPro 64
Qy 1521 GAAGCTGAAGAGACTTGTTCACAAAGTACACAGTGGGAACACTTACCGAGTGGGTGACACT 1580
Db 65 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 84
Qy 1581 TATGAGCGTCTTAAGACTCCATGATCTGGGACTGTACCTGCATCGGGCTGGGCGAGGG 1640
Db 85 TyrGluArgProLysAspSerMetIleTyrAspCysThrCysIleGlyAlaGlyArgGly 104
Qy 1641 AGAATAAGCTGTACCATC 1658
```

```
Db 105 ArgileSerCysThrIle 110
|||||
RESULT 14
US-10-680-287A-677
; Sequence 677, Application US/10680287A
; Publication No. US20040244069A1
; GENERAL INFORMATION:
; APPLICANT: Babij, Philip
; APPLICANT: Yaworsky, Paul
; APPLICANT: Bex, Frederick J. III
; APPLICANT: Bodine, Peter Van Nest
; TITLE OF INVENTION: Transgenic Animal Model of Bone Mass Modulation
; FILE REFERENCE: 032796-179
; CURRENT APPLICATION NUMBER: US/10/680,287A
; CURRENT FILING DATE: 2003-10-08
; PRIOR APPLICATION NUMBER: PCT/US02/14876
; PRIOR FILING DATE: 2002-05-13
; PRIOR APPLICATION NUMBER: US 60/290,071
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/291,311
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: US 60/353,058
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: US 60/361,293
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 812
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 677
; LENGTH: 2328
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-680-287A-677

Alignment Scores:
Pred. No.: 1.3e-47 Length: 2328
Score: 600.00 Matches: 106
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.42% Indels: 0
DB: 5 Gaps: 0

US-09-940-235-10 (1-1661) x US-10-680-287A-677 (1-2328)
QY 1341 CAGGCGCAGCAATGTTTTCAGCCCCAGTCCCGGTGGCTGTCTCAGTCAAAAGCAAGCCCGGT 1400
Db 5 GlnAlaGlnGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 24
QY 1401 TGTATTGACAATGGAACACACTATCAGATAAATCAACAGTGGGAGCGGACCTACCTAGGT 1460
Db 25 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 44
QY 1461 AATGTGTGTGTTTGTACTTGTATTGAGGAGGAGCGAGGTTTAACTCGGAAAGTAAACCT 1520
Db 45 AsnValLeuValCysThrCysTyrGlySerArgGlyPheAsnGlyPheAsnGlySerLysPro 64
QY 1521 GAAGCTGAAGAGACTTGTCTTTGACAAGTACACTGGGAAACACTTACCGAGTGGGTGACACT 1580
Db 65 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 84
QY 1581 TATGAGCGCTTAAGACTCCATGATCTGGGACTGTACCTGCTGCTGCGGCTGGCGGAGG 1640
Db 85 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 104
QY 1641 AGAATAAGCTGTACCATC 1658
Db 105 ArgileSerCysThrIle 110

Search completed: January 28, 2006, 02:56:50
Job time : 128.622 secs

Db 105 ArgileSerCysThrIle 110
|||||
RESULT 15
US-10-477-173-677
; Sequence 677, Application US/10477173
; Publication No. US20050070699A1
; GENERAL INFORMATION:
; APPLICANT: Genome Therapeutics Corporation and
; APPLICANT: Allen, Kristina M.
; APPLICANT: Yaworsky, Paul
; APPLICANT: Morales, Arturo J.
; APPLICANT: Graham, James R.
; APPLICANT: Anisowicz, Anthony
; APPLICANT: Liu, Wei
; TITLE OF INVENTION: HBM Variants that Modulate Bone Mass and Lipid Levels
; FILE REFERENCE: 032796-135
; CURRENT APPLICATION NUMBER: US/10/477,173
; CURRENT FILING DATE: 2003-11-10
; PRIOR APPLICATION NUMBER: US 60/290,071
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/291,311
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: US 60/353,058
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: US 60/361,293
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 1086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 677
; LENGTH: 2328
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-477-173-677

Alignment Scores:
Pred. No.: 1.3e-47 Length: 2328
Score: 600.00 Matches: 106
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.42% Indels: 0
DB: 5 Gaps: 0

US-09-940-235-10 (1-1661) x US-10-477-173-677 (1-2328)
QY 1341 CAGGCGCAGCAATGTTTTCAGCCCCAGTCCCGGTGGCTGTCTCAGTCAAAAGCAAGCCCGGT 1400
Db 5 GlnAlaGlnGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 24
QY 1401 TGTATTGACAATGGAACACACTATCAGATAAATCAACAGTGGGAGCGGACCTACCTAGGT 1460
Db 25 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 44
QY 1461 AATGTGTGTGTTTGTACTTGTATTGAGGAGGAGCGAGGTTTAACTCGGAAAGTAAACCT 1520
Db 45 AsnValLeuValCysThrCysTyrGlySerArgGlyPheAsnGlyPheAsnGlySerLysPro 64
QY 1521 GAAGCTGAAGAGACTTGTCTTTGACAAGTACACTGGGAAACACTTACCGAGTGGGTGACACT 1580
Db 65 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 84
QY 1581 TATGAGCGCTTAAGACTCCATGATCTGGGACTGTACCTGCTGCTGCGGCTGGCGGAGG 1640
Db 85 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 104
QY 1641 AGAATAAGCTGTACCATC 1658
Db 105 ArgileSerCysThrIle 110

Search completed: January 28, 2006, 02:56:50
Job time : 128.622 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 28, 2006, 02:09:11 ; Search time 4.80939 Seconds
(without alignments)
7479.985 Million cell updates/sec

Title: US-09-940-235-10
Perfect score: 2939
Sequence: 1 gcaacccgccagctagcc.....gaataagctgtaccattctaa 1661

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 75621 seqs, 10829074 residues

Total number of hits satisfying chosen parameters: 151242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09940235/runat_27012006_144220_27676/app_query.fasta_1.7708
-DB=Published Applications AA New -QFMT=faстан -SUFFIX=rapbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09940235 @CGN_1_1_1@runat_27012006_144220_27676
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORE=0 -WAIT -DSPBLOCK=100
-LONGLOG -DRV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-XGAPOP=6 -XGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA New.*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pcp.*
2: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pcp.*
3: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pcp.*
4: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pcp.*
5: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pcp.*
6: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pcp.*
7: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pcp.*
8: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1984	67.5	414	6	US-10-631-558-2
2	1928	65.6	413	7	US-11-032-951-12
3	600	20.4	259	6	US-10-631-558-4
4	596	20.3	300	7	US-11-006-119-31
5	596	20.3	642	6	US-10-995-561-631
6	596	20.3	657	6	US-10-995-561-622
7	596	20.3	984	6	US-10-995-561-629
8	596	20.3	2296	6	US-10-995-561-633
9	596	20.3	2355	6	US-10-995-561-623
10	596	20.3	2355	6	US-10-995-561-627

11	596	20.3	2384	6	US-10-821-234-1545
12	596	20.3	2386	6	US-10-995-561-626
13	214	7.3	1341	6	US-10-995-561-621
14	212	7.2	1259	6	US-10-995-561-625
15	211.5	7.2	1286	6	US-10-995-561-628
16	211	7.2	1348	6	US-10-995-561-624
17	208	7.1	693	6	US-10-995-561-632
18	208	7.1	1315	6	US-10-995-561-630
19	199	6.8	847	6	US-10-995-561-634
20	139	4.7	24	7	US-11-006-119-30
21	131.5	4.5	2047	7	US-11-013-759-4
22	131.5	4.5	2047	7	US-11-013-759-7
23	130.5	4.4	1992	7	US-11-013-759-3
24	130.5	4.4	1992	7	US-11-013-759-13
25	119	4.0	801	7	US-11-174-150-29
26	119	4.0	801	7	US-11-124-368A-292
27	119	4.0	1417	7	US-11-052-554A-8
28	117	4.0	2314	7	US-11-013-759-11
29	116.5	4.0	824	6	US-10-821-234-1008
30	115	3.9	21	7	US-11-032-951-7
31	113	3.8	470	6	US-10-485-517-319
32	111	3.8	407	7	US-11-016-564-7
33	110.5	3.8	1765	6	US-10-055-877-140
34	110	3.7	1940	6	US-10-055-877-141
35	109	3.7	2890	7	US-11-115-639-31
36	109	3.7	2890	7	US-11-115-639-32
37	109	3.7	2890	7	US-11-115-639-33
38	108.5	3.7	463	6	US-10-733-626-960
39	108	3.7	1647	7	US-11-052-554A-260
40	107.5	3.7	2515	7	US-11-113-424-53
41	107	3.6	2710	7	US-11-051-453-41
42	106.5	3.6	1193	7	US-11-115-639-19
43	106.5	3.6	1193	7	US-11-115-639-20
44	106.5	3.6	1193	7	US-11-115-639-21
45	106	3.6	876	7	US-11-077-550-82

ALIGNMENTS

RESULT 1
US-10-631-558-2
; Sequence 2, Application US/10631558
; Publication No. US20050260598A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammar
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir

; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE

; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID

; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/10/631,558
; CURRENT FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US/09/940,235
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FAST-SEQ for Windows Version 4.0
; SEQ ID NO 2

; LENGTH: 414
; TYPE: PRT
; ORGANISM: Streptococcus equisimilis
US-10-631-558-2

Alignment Scores:

```

Pred. No.: 9,7e-141 Length: 414
Score: 1984.00 Matches: 383
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 67.51% Indels: 0
DB: 6 Gaps: 0

US-09-940-235-10 (1-1661) x US-10-631-558-2 (1-414)

QY 183 ATTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAAATGGTTGTT 242
Db 1 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal 20
QY 243 AGCGTTGCTGGTACTGTTGAGGCGACGAATCAAGACATTAGTCTTAAATTTTTTGAATC 302
Db 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluLe 40
QY 303 GATCTAACATCAGCACCTGCTCATGGAGGAAAGACAGACGAGCTTAAAGTCCAAATCA 362
Db 41 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 60
QY 363 AAACCATTTGCTACTGATAGTGGCGGATGTCATAAACTTGAGAAAGCTGACTACTA 422
Db 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLe 80
QY 423 AAGGCTATTCAAGAAACAATTGATCGCTAAACGTCACAGTAAACGACGACTACTTTGAGTC 482
Db 81 LysAlaIleGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100
QY 483 ATTGATTTTGCAAGCGATGCAACCATTTACTGTATCGAAACGCGCAAGGTCTACTTTGCTGAC 542
Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
QY 543 AAAGATGTTGCGTAACTGTCGCGACCCAACTGTCCTCAAGAAATTTTGTCTAAGCGACAT 602
Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
QY 603 GTGCGGTTAGCACCATATAAGAAACCAATACAAACCAACGCGAAATCTGTTGATGTC 662
Db 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
QY 663 GAATATACCTGACAGTTTACTCCCTTAAACCTGTAGTACGATTTTCAGACAGGTCTCAAA 722
Db 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
QY 723 GATACTAAGCTATTGAAACACTAGCTATPCGGTGCACCATCACATCTCAAGAATTTACTA 782
Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 200
QY 783 GCTCAAGCACAAAGCATTTTAAACAAACACCCAGGCTATACGATTTTATGAACTGAC 842
Db 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 220
QY 843 TCCTCAATCGTCACATCATGACAAATGACATTTTCGTTACGATTTTACCAATGGATCAAGAG 902
Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
QY 903 TTTACTTACCGTGTAAAAATCGGAAACAAAGCTTATAGGATCAATAAAAAATCTGCTG 962
Db 241 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 260
QY 963 AATGAAGAAATAAACAACTGACCTGATCTCTGAGAAATATATAGCTCTTAAAAAGGG 1022
Db 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280
QY 1023 GAAAGCGGTATGATCCCTTTTCATCGCTGCTGAACTGTTCACCATCAAAATCGTT 1082
Db 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
QY 1083 GATGTCGATACCAACGAATTTGCTAAAAAGTGGAGCAGCTCTTAAACAGCTAGCGCAACGTAAC 1142
Db 301 AspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320

```

RESULT 2

```

US-11-032-951-12
; Sequence 12, Application US/11032951
; Publication No. US20050282741A1
; GENERAL INFORMATION:
; APPLICANT: Krystal, Gerald
; APPLICANT: Rabkin, Simon W.
; TITLE OF INVENTION: Peptides and Their Use to Ameliorate
; TITLE OF INVENTION: Cell Death
; FILE REFERENCE: 50216/003005
; CURRENT APPLICATION NUMBER: US/11/032,951
; PRIOR FILING DATE: 2005-01-11
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 09/919,703
; PRIOR FILING DATE: 1999-04-19
; PRIOR APPLICATION NUMBER: US 09/294,457
; PRIOR FILING DATE: 1996-12-05
; PRIOR APPLICATION NUMBER: US 60/008,233
; PRIOR FILING DATE: 1995-12-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 413
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic polypeptide
US-11-032-951-12

```

```

Alignment Scores:
Pred. No.: 1,43e-136 Length: 413
Score: 1928.00 Matches: 377
Percent Similarity: 98.95% Conservative: 1
Best Local Similarity: 98.69% Mismatches: 2
Query Match: 65.60% Indels: 2
DB: 7 Gaps: 2

```

US-09-940-235-10 (1-1661) x US-11-032-951-12 (1-413)

```

QY 183 ATTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAAATGGTTGTT 242
Db 1 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal 20
QY 243 AGCGTTGCTGGTACTGTTGAGGCGACGAATCAAGACATTAGTCTTAAATTTTTTGAATC 302
Db 21 SerValAlaGlyThrVal---GlyThrAsnGlnAspIleSerLeuLysPhePheGluLe 39
QY 303 GATCTAACATCAGCACCTGCTCATGGAGGAAAGACAGACGAGCTTAAAGTCCAAATCA 362
Db 40 AspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 59
QY 363 AAACCATTTGCTACTGATAGTGGCGGATGTCATAAACTTGAGAAAGCTGACTTACTA 422
Db 60 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLe 79
QY 423 AAGGCTATTCAAGAAACAATTGATCGCTTAAACGTCACAGTAAACGACGACTACTTTGAGTC 482

```



```

Db      80  LysAlaIleGlnGluLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 99
QY      483  ATTGATTTCGAAGCGATGCAACCATCTACTGTCGAACCGCAGGCTCTACTTTGTCGAC 542
Db      100  IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyIysValTyrPheAlaAsp 119
QY      543  AAGATGTTGCGTAACCTTTCGCGACCCCAACCTGTCTCAAGAATTTTTCGTAAGCGGACAT 602
Db      120  LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 139
QY      603  CTCGCGGTAGACCATATAAAGAAAAACCAATACAAACCAAGCGAAATCTGTGTGATGTG 662
Db      140  ValArgValArgProTyrIysGluIysProIleGlnAsnGlnAlaIysSerValAspVal 159
QY      663  GAATATACCTGTCAGATTCTCTCCCTTAAACCCCTGATGACGACCTTCAGACGAGCTCTCAA 722
Db      160  GluTyrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuLys 179
QY      723  GATCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAACTACTA 782
Db      180  LeuThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 199
QY      783  GCTCAAGCAAAAGCAATTTTAAACAAACCCAGCGCTATACGATTTATGACGTCGAC 842
Db      200  AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 219
QY      843  TCCTCAATCGTCTCATGACATGACATTTTCGTACGATTTTACCAATGGATCAAGAG 902
Db      220  SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 239
QY      903  TTTACTTACCGTGTAAATAATCGGCAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 962
Db      240  PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysSerGlyLeu 259
QY      963  AATGAAGAAATAACACACTGACCTGATCTCTCGAAGAAATAT---TACGTCCTTTAAAAA 1019
Db      260  AsnGluGluIleAsnAsnThrAspLeuIleSerLeuGluTyrLysTyrValLeuLysLys 279
QY      1020  GGGGAAAGCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTACCAATCAATAC 1079
Db      280  GlyGluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyr 299
QY      1080  GTTGATGTCGATACCAACGAATTCCTAAAAGTGACGAGCTCTTAACAGCTAGCGACCT 1139
Db      300  ValAspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArg 319
QY      1140  AACTTAGACTTCAGAGATTTATACGATCTCTCGTGATAAGGCTAAACTTACTCTACAACAAT 1199
Db      320  AsnLeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsn 339
QY      1200  CTCGATGCTTTTGGTATTATGACTATACCTTAAGTGAAGAGTAGAGGATAATACGAT 1259
Db      340  LeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAsp 359
QY      1260  GACACCAACCGTATCATACCGTTTATATGGCAAGCGACCGAGAGAGAGATGCTAGC 1319
Db      360  AspThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSer 379
QY      1320  TATCAT 1325
Db      380  TyrHis 381

```

RESULT 3

```

US-10-631-558-4
; Sequence 4, Application US/10631558
; Publication No. US20050260598A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammarra
; APPLICANT: Nihalani, Deepak

```

```

; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/10/631,558
; CURRENT FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US/09/940,235
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-631-558-4

Alignment Scores:
Pred. No.:          9,22e-38          Length:          259
Score:             600.00             Matches:         106
Percent Similarity: 100.00%           Conservative:    0
Best Local Similarity: 100.00%         Mismatches:     0
Query Match:       20.42%             Indels:         0
DB:                6                  Gaps:          0

US-09-940-235-10 (1-1661) x US-10-631-558-4 (1-259)

QY      1341  CAGCGCAGCAATGGTTTCAGCCCACTCCCGGTGGCTGTCACTCAAGCAAGCCCGGT 1400
Db      1   GlnAlaGlnGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 20
QY      1401  TGTATCATCATGGAACACACTATCATATAAACAACAGTGGGAGCGGACCTACCTAGGT 1460
Db      21  CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 40
QY      1461  AATGTGTGTGTGTGTACTTGTATGTGAGAGAACCGAGGTTTAACTCGGAAAGTAAACCT 1520
Db      41  AsnValLeuValCysThrCysTyrGlySerArgGlyPheAsnCysGluSerLysPro 60
QY      1521  GAAGCTGAAGAGACTTGTCTTTTGACACACTGGAACACTTACCGAGTGGGTGACACT 1580
Db      61  GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
QY      1581  TATGAGCGTCTTAAAGACTCATGATCTGGGACTGTACCTGATCGGGGCTGGCGAGGG 1640
Db      81  TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 100
QY      1641  AGAATAAGCTGTACCATC 1658
Db      101  ArgIleSerCysThrIle 106

RESULT 4
US-11-006-119-31
; Sequence 31, Application US/11006119
; Publication No. US20050260691A1
; GENERAL INFORMATION:
; APPLICANT: Ward, Brian
; APPLICANT: Ndao, Momar
; APPLICANT: Caffrey, Rebecca
; APPLICANT: Spithill, Terry
; APPLICANT: Li, Hongshan
; APPLICANT: Podust, Vladimir
; APPLICANT: Perichon, Regis
; APPLICANT: CIPHERGEN Biosystems, Inc.
; TITLE OF INVENTION: Serum Biomarkers for Chagas Disease
; FILE REFERENCE: 016866-012130US
; CURRENT APPLICATION NUMBER: US/11/006,119

```

;
; CURRENT FILING DATE: 2004-12-06
; PRIOR APPLICATION NUMBER: US 60/527,153
; PRIOR FILING DATE: 2003-12-05
; PRIOR APPLICATION NUMBER: US 60/565,093
; PRIOR FILING DATE: 2004-04-22
; PRIOR APPLICATION NUMBER: US 60/625,519
; PRIOR FILING DATE: 2004-11-06
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 31
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: N-terminal 28.7 fragment of Fibronectin
US-11-006-119-31

Alignment Scores:
Pred. No.: 1,86e-37 Length: 300
Score: 596.00 Matches: 105
Percent Similarity: 99.06% Conservative: 0
Best Local Similarity: 99.06% Mismatches: 1
Query Match: 20.28% Indels: 0
DB: Gaps: 0

US-09-940-235-10 (1-1661) x US-11-006-119-31 (1-300)

QY 1341 CAGGCGCAGCAATGGTTGAGCCCGAGTCCCGGTGGCTGTCAGTCAAGCAAGCCGGT 1400
Db 1 GlnAlaGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 20
QY 1401 TGTATTGACAAATGGAACACTATCAGATAAATCAACAGTGGGAGCGACCTACCTAGT 1460
Db 21 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnInTrpGluArgThrTyrLeuGly 40
QY 1461 AATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1520
Db 41 AsnAlaLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGluSerLysPro 60
QY 1521 GAAGCTGAAGAGACTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1580
Db 61 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
QY 1581 TATGAGCGTCCTAAAGACTCCATGATCTGGAGTGTACCTGCGGCTGGGCTGGGCGAGG 1640
Db 81 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 100
QY 1641 AGAATAAGCTGTACCATC 1658
Db 101 ArgIleSerCysThrIle 106

RESULT 5
US-10-995-561-631
; Sequence 631, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 631
; LENGTH: 642
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-631

Alignment Scores:
Pred. No.: 2.03e-37 Length: 642

Score: 596.00 Matches: 105
Percent Similarity: 99.06% Conservative: 0
Best Local Similarity: 99.06% Mismatches: 1
Query Match: 20.28% Indels: 0
DB: Gaps: 0

US-09-940-235-10 (1-1661) x US-10-995-561-631 (1-642)

QY 1341 CAGGCGCAGCAATGGTTGAGCCCGAGTCCCGGTGGCTGTCAGTCAAGCAAGCCGGT 1400
Db 32 GlnAlaGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 51
QY 1401 TGTATTGACAAATGGAACACTATCAGATAAATCAACAGTGGGAGCGACCTACCTAGT 1460
Db 52 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnInTrpGluArgThrTyrLeuGly 71
QY 1461 AATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1520
Db 72 AsnAlaLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGluSerLysPro 91
QY 1521 GAAGCTGAAGAGACTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1580
Db 92 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 111
QY 1581 TATGAGCGTCCTAAAGACTCCATGATCTGGAGTGTACCTGCGGCTGGGCTGGGCGAGG 1640
Db 112 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 131
QY 1641 AGAATAAGCTGTACCATC 1658
Db 132 ArgIleSerCysThrIle 137

RESULT 6

US-10-995-561-622
; Sequence 622, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 622
; LENGTH: 657
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-622

Alignment Scores:
Pred. No.: 2.04e-37 Length: 657
Score: 596.00 Matches: 105
Percent Similarity: 99.06% Conservative: 0
Best Local Similarity: 99.06% Mismatches: 1
Query Match: 20.28% Indels: 0
DB: Gaps: 0

US-09-940-235-10 (1-1661) x US-10-995-561-622 (1-657)

QY 1341 CAGGCGCAGCAATGGTTGAGCCCGAGTCCCGGTGGCTGTCAGTCAAGCAAGCCGGT 1400
Db 32 GlnAlaGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 51
QY 1401 TGTATTGACAAATGGAACACTATCAGATAAATCAACAGTGGGAGCGACCTACCTAGT 1460
Db 52 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnInTrpGluArgThrTyrLeuGly 71
QY 1461 AATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1520
Db 72 AsnAlaLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGluSerLysPro 91

QY 1521 GAAGCTGAAGAGACTTGGCTTTGACAAAGTACACTGGGAACACTTACCGAGTGGGTGACACT 1580
 |||||
 Db 92 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 111
 |||||
 QY 1581 TATGAGCGTCTTAAGACTCCATGATCTGGGACTGTACCTGCATCGGGGCTGGCGAGGG 1640
 |||||
 Db 112 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 131
 |||||
 QY 1641 AGAATAAGCTGTACCAC 1658
 |||||
 Db 132 ArgIleSerCysThrIle 137
 |||||

RESULT 7

US-10-995-561-629
 ; Sequence 629, Application US/10995561
 ; Publication No. US20050272054A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele et al.
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
 ; FILE REFERENCE: CL001559
 ; CURRENT APPLICATION NUMBER: US/10/995,561
 ; CURRENT FILING DATE: 2004-11-24
 ; NUMBER OF SEQ ID NOS: 85702
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 629
 ; LENGTH: 984
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-995-561-629

Alignment Scores:
 Pred. No.: 2,148-37 Length: 984
 Score: 596.00 Matches: 105
 Percent Similarity: 99.06% Conservative: 0
 Best Local Similarity: 99.06% Mismatches: 1
 Query Match: 20.28% Indels: 0
 DB: 6 Gaps: 0

US-09-940-235-10 (1-1661) x US-10-995-561-629 (1-984)

QY 1341 CAGCGCAGCAAAATGGTTTCAGCCCGTCCCGTGTCTGCTCAGTCAAAAGCAAGCCCGGT 1400
 |||||
 Db 32 GlnAlaGlnGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 51
 |||||
 QY 1401 TGTATGACAATCGAAACACTATCAGATAAATCAACAGTGGGAGCGGACCTACCTAGGT 1460
 |||||
 Db 52 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnTrpGluArgThrTyrLeuGly 71
 |||||
 QY 1461 AATGTTGGTTGTACTTGTATGAGGAGCGGAGGTTTAACTCGGAAGTAAACCT 1520
 |||||
 Db 72 AsnAlaLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnGlySerLysPro 91
 |||||
 QY 1521 GAAGCTGAAGAGACTTGGCTTTGACAAAGTACACTGGGAGCGGTTTAACTCGGAAGTAAACCT 1580
 |||||
 Db 92 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 111
 |||||
 QY 1581 TATGAGCGTCTTAAGACTCCATGATCTGGGACTGTACCTGCATCGGGGCTGGCGAGGG 1640
 |||||
 Db 112 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 131
 |||||
 QY 1641 AGAATAAGCTGTACCAC 1658
 |||||
 Db 132 ArgIleSerCysThrIle 137
 |||||

RESULT 8

US-10-995-561-633
 ; Sequence 633, Application US/10995561
 ; Publication No. US20050272054A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele et al.

; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
 ; FILE REFERENCE: CL001559
 ; CURRENT APPLICATION NUMBER: US/10/995,561
 ; CURRENT FILING DATE: 2004-11-24
 ; NUMBER OF SEQ ID NOS: 85702
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 633
 ; LENGTH: 2296
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-995-561-633

Alignment Scores:
 Pred. No.: 2,368-37 Length: 2296
 Score: 596.00 Matches: 105
 Percent Similarity: 99.06% Conservative: 0
 Best Local Similarity: 99.06% Mismatches: 1
 Query Match: 20.28% Indels: 0
 DB: 6 Gaps: 0

US-09-940-235-10 (1-1661) x US-10-995-561-633 (1-2296)

QY 1341 CAGCGCAGCAAAATGGTTTCAGCCCGTCCCGTGTCTGCTCAGTCAAAAGCAAGCCCGGT 1400
 |||||
 Db 32 GlnAlaGlnGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 51
 |||||
 QY 1401 TGTATGACAATCGAAACACTATCAGATAAATCAACAGTGGGAGCGGACCTACCTAGGT 1460
 |||||
 Db 52 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnTrpGluArgThrTyrLeuGly 71
 |||||
 QY 1461 AATGTTGGTTGTACTTGTATGAGGAGCGGAGGTTTAACTCGGAAGTAAACCT 1520
 |||||
 Db 72 AsnAlaLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnGlySerLysPro 91
 |||||
 QY 1521 GAAGCTGAAGAGACTTGGCTTTGACAAAGTACACTGGGAGCGGTTTAACTCGGAGTGGGTGACACT 1580
 |||||
 Db 92 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 111
 |||||
 QY 1581 TATGAGCGTCTTAAGACTCCATGATCTGGGACTGTACCTGCATCGGGGCTGGCGAGGG 1640
 |||||
 Db 112 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 131
 |||||
 QY 1641 AGAATAAGCTGTACCAC 1658
 |||||
 Db 132 ArgIleSerCysThrIle 137
 |||||

RESULT 9

US-10-995-561-623
 ; Sequence 623, Application US/10995561
 ; Publication No. US20050272054A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele et al.
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
 ; FILE REFERENCE: CL001559
 ; CURRENT APPLICATION NUMBER: US/10/995,561
 ; CURRENT FILING DATE: 2004-11-24
 ; NUMBER OF SEQ ID NOS: 85702
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 623
 ; LENGTH: 2355
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-995-561-623

Alignment Scores:

Pred. No.: 2,368-37 Length: 2355
 Score: 596.00 Matches: 105
 Percent Similarity: 99.06% Conservative: 0
 Best Local Similarity: 99.06% Mismatches: 1

```
Query Match: 20.28% Indels: 0
DB: 6 Gaps: 0
US-09-940-235-10 (1-1661) x US-10-995-561-623 (1-2355)
QY 1341 CAGGCGCAGCAATGTTTCAGCCCGAGTCCCGGTGGCTGTCAAGTCAAGCAAGCCCGGT 1400
DB 32 GlnAlaGlnGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 51
QY 1401 TGTATTGACAATGGAACACATCATCATGATAAATCAACAGTGGAGCGGACCTACCTAGGT 1460
DB 52 CysTyrAspAsnGlyHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 71
QY 1461 AATGTTGTTGTTGTTACTTGTATGAGGAAGCCGAGGTTTAACTGCGAAAGTAAACCT 1520
DB 72 AsnAlaLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGluSerLysPro 91
QY 1521 GAAGCTGAAGAGACTGTTTGCACAAAGTACACTGGGAACACTTACCAGTGGGTGACACT 1580
DB 92 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 111
QY 1581 TATGAGCGTCTAAAGACTCCATGATCTGGAGTGTACCTGCATCGGGGTGGCGAGGG 1640
DB 112 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 131
QY 1641 AGAATAAGCTGTACCATC 1658
DB 132 ArgIleSerCysThrIle 137
RESULT 10
US-10-995-561-627
; Sequence 627, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 627
; LENGTH: 2355
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-627
Alignment Scores:
Pred. No.: 2,368-37 Length: 2355
Score: 596.00 Matches: 105
Percent Similarity: 99.06% Conservative: 0
Best Local Similarity: 99.06% Mismatches: 1
Query Match: 20.28% Indels: 0
DB: 6 Gaps: 0
US-09-940-235-10 (1-1661) x US-10-995-561-627 (1-2355)
QY 1341 CAGGCGCAGCAATGTTTCAGCCCGAGTCCCGGTGGCTGTCAAGTCAAGCAAGCCCGGT 1400
DB 32 GlnAlaGlnGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 51
QY 1401 TGTATTGACAATGGAACACATCATCATGATAAATCAACAGTGGAGCGGACCTACCTAGGT 1460
DB 52 CysTyrAspAsnGlyHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 71
QY 1461 AATGTTGTTGTTGTTACTTGTATGAGGAAGCCGAGGTTTAACTGCGAAAGTAAACCT 1520
DB 72 AsnAlaLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGluSerLysPro 91
QY 1521 GAAGCTGAAGAGACTGTTTGCACAAAGTACACTGGGAACACTTACCAGTGGGTGACACT 1580
DB 92 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 111
QY 1581 TATGAGCGTCTAAAGACTCCATGATCTGGAGTGTACCTGCATCGGGGTGGCGAGGG 1640
DB 112 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 131
QY 1641 AGAATAAGCTGTACCATC 1658
DB 132 ArgIleSerCysThrIle 137
RESULT 11
US-10-821-234-1545
; Sequence 1545, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: PC SEQ_genes Version 1.0
; SEQ ID NO 1545
; LENGTH: 2384
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1545
Alignment Scores:
Pred. No.: 2,378-37 Length: 2384
Score: 596.00 Matches: 105
Percent Similarity: 99.06% Conservative: 0
Best Local Similarity: 99.06% Mismatches: 1
Query Match: 20.28% Indels: 0
DB: 6 Gaps: 0
US-09-940-235-10 (1-1661) x US-10-821-234-1545 (1-2384)
QY 1341 CAGGCGCAGCAATGTTTCAGCCCGAGTCCCGGTGGCTGTCAAGTCAAGCAAGCCCGGT 1400
DB 61 GlnAlaGlnGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 80
QY 1401 TGTATTGACAATGGAACACATCATCATGATAAATCAACAGTGGAGCGGACCTACCTAGGT 1460
DB 81 CysTyrAspAsnGlyHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 100
QY 1461 AATGTTGTTGTTGTTACTTGTATGAGGAAGCCGAGGTTTAACTGCGAAAGTAAACCT 1520
DB 101 AsnAlaLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGluSerLysPro 120
QY 1521 GAAGCTGAAGAGACTGTTTGCACAAAGTACACTGGGAACACTTACCAGTGGGTGACACT 1580
DB 121 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 140
QY 1581 TATGAGCGTCTAAAGACTCCATGATCTGGAGTGTACCTGCATCGGGGTGGCGAGGG 1640
DB 141 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 160
QY 1641 AGAATAAGCTGTACCATC 1658
DB 161 ArgIleSerCysThrIle 166
RESULT 12
US-10-995-561-626
; Sequence 626, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
```

; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: DETECTION AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 626
; LENGTH: 2386
; TYPE: PRN
; ORGANISM: Homo sapiens
US-10-995-561-626

Alignment Scores:
Pred. No.: 2,37e-37 Length: 2386
Score: 596.00 Matches: 105
Percent Similarity: 99.06% Conservative: 0
Best Local Similarity: 99.06% Mismatches: 1
Query Match: 20.28% Indels: 0
DB: Gaps: 0

US-09-940-235-10 (1-1661) x US-10-995-561-626 (1-2386)

QY 1341 CAGGCGCAGCAATGTTGAGCCCGAGTCCCGTGGCTGCTCAGTCAAGCAAGCCCGT 1400
DB 32 GlnAlaGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 51
QY 1401 TGTATGACAATGGAAACATATCAGATTAATCAACAGTGGAGCGGACTACCTAGGT 1460
DB 52 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnInTrpGluArgThrTyrLeuGly 71
QY 1461 AATGTGTTGTTGTTACTTGTATGAGGAGCGAGGTTTAACTCGGAAAGTAAACCT 1520
DB 72 AsnAlaLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGluSerLysPro 91
QY 1521 GAAGCTGAAGAGACTTGTCTTGAACAGTACACTGGGAACACTTACCGAGTGGGTGACACT 1580
DB 92 GluAlaGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 111
QY 1581 TATGAGCTTCAAGACTCCATGATCTGGAGTCTGACTGATCGATCGGGGTGGCGAGGG 1640
DB 112 TyrGluArgProLysAspSerMetIleTyrAspCysThrCysIleGlyAlaGlyArgGly 131
QY 1641 AGAATAAGCTGTACCATC 1658
DB 132 ArgIleSerCysThrIle 137

RESULT 13

US-10-995-561-621
; Sequence 621, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 621
; LENGTH: 1341
; TYPE: PRN
; ORGANISM: Homo sapiens
US-10-995-561-621

Alignment Scores:
Pred. No.: 5.95e-09 Length: 1341
Score: 214.00 Matches: 133
Percent Similarity: 32.70% Conservative: 73
Best Local Similarity: 21.11% Mismatches: 200

Query Match: 7.28% Indels: 224
DB: Gaps: 29
US-09-940-235-10 (1-1661) x US-10-995-561-621 (1-1341)
QY 15 CTAGCCGGTCTCTCAACGACGAGGACGATCATCGCGCGCGGCGCAGACCCCAAGCGC 74
DB 789 LeuIleGlyThrGlnSerThrAlaIleProAlaProThrAspLeuLysPheThrGlnVal 808
QY 75 TGCCCGAGATCTCGATCCCGGAA-----ATTAAATACCACTCATATAGG 119
DB 809 ThrProThrSerLeuSerAlaGlnInTrpThrProProAsnValGlnLeuThrGlyTyrArg 828
QY 120 GAGACCACACGCGTTTCCCTCTAGAAATAATTTTGTAACTTTAAGAAGGAGATATACC 179
DB 829 ValArgValThrProLys----- 834
QY 180 ATGATTGCTGACCTCGAGTGGCTGTAGACGTCATCTGCAACACAGCAATGTGTT 239
DB 835 GluLysThrGlyProMetLysGluIleAsnLeuAlaProAspSerSerSer----- 851
QY 240 GTTAGCGTGTCTGCTACTGTTGAGGGGACGAATCAAGACATAGTCTTAAATTTTTCGA 299
DB 852 ValValValSerGlyLeuMetValAlaThrLysTyrGluValSerValTyrAlaLeuLys 871
QY 300 ATCGATCTAAACATCAGCACCTGCTCATGGAGGAAAGACAGAG---CAAGGCTTAAGTCCA 356
DB 872 AspThrLeuThrSerArgProAlaGlnGlyValValThrThrLeuGluAsnValSerPro 891
QY 357 AAATCAAAACCATTTGCTACTGATAGTGGCGCGATGTACATAAATTTGAGAAAGCTGAC 416
DB 892 ProArgArgAlaArgValThrAsp----- 899
QY 417 TTACTAAGGCTATTCAAGAACAAATTGATCGCTAACGTCACAGTAACGACGACTACTTT 476
DB 899 ----- 899
QY 477 GAGTCTATTGATTTTGAAGCGATGCAACCACTTACTGATCGAAGCGCAAGGTCTACTTT 536
DB 900 -----AlaThrGluThrThrIleThrIle----- 907
QY 537 GCTGACAAAGATGTTTCGTAACCTTGCAGCAACCACTGTCCTCAAGAAATTTTGTCAAGC 596
DB 908 -----SerTyrArgThrLysThrGluThrIleThrGlyPheGlnValAsp 922
QY 597 GGACATGTGCGGTTAGACCATATAAGAAAAACAATACAAACCAAGCAAGCAATCTGTT 656
DB 923 Ala-----ValProAlaAsnGlyGlnThrProIleGlnArgThrIleLysProAsp 939
QY 657 GATGTGGAATATACTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTCAAGACGAGT 716
DB 940 ValArgSerTyrThrIle-----ThrGlyLeuGlnProGlyThrAspTyrLys----- 955
QY 717 CTCAAAGATATAAGCTATTGAAACACACTAGCTATCGGTGACACCATCACATCTCAAGAA 776
DB 956 -----IleTyrLeuTyrThrLeuAsnAspAsnAlaArgSerSerPro 969
QY 777 TTACTAGCTCAAGCA----- 791
DB 970 ValValIleAspAlaSerThrAlaIleAspAlaProSerAsnLeuArgPheLeuAlaThr 989
QY 792 -----CAAAAGCATTTTAAACAAAACCAACCCA-----GGCTATACG 827
DB 990 ThrProAsnSerLeuLeuValSerTrpGlnProProArgAlaArgIleThrGlyTyrIle 1009
QY 828 ATT---TATGAAGTGTACTCTCAATCGTCACTCATGCAATGACATTTTTCGTCAGGATT 884
DB 1010 IleLysTyrGluLysProGlySerProPro-----ArgGluVal 1022
QY 885 TTACCAATGGATCAA----- 899
DB 1023 ValProArgProArgProGlyValThrGluAlaThrIleThrGlyLeuGluProGlyThr 1042


```

QY 792 -----CAAGCATTTTAAACAAACCACCA-----GGCTATACG 827
Db 990 ThrProAsnSerLeuLeuValSerTrpGlnProProArgAlaArgIleThrGlyTyrIle 1009
QY 828 ATT---TATGAAGTCTCTCAATCGTCACTCATGACAATGACATTTTCCGACGATT 884
Db 1010 IleLysTyrGluLysProGlySerProPro-----ArgGluVal 1022
QY 885 TTACCAATGGATCA----- 899
Db 1023 ValProArgProArgProGlyValThrGluAlaThrIleThrGlyLeuGluProGlyThr 1042
QY 900 GAGTTTACTTACCGTGT-----AAAAATCGGAACAAGCTTATAGATCAAT 947
Db 1043 GluTyrThrIleTyrValIleAlaLeuLysAsnGlnLysSerGluProLeuIleGly 1062
QY 948 AAAAATCTGGTCTGAATGAAGAATAAACAAC---ACTGACCTGATCTCTGAGAATAT 1004
Db 1063 ArgLysLysThrValAsnGluGlyLeuAsnGlnProThrAsp----- 1076
QY 1005 TACGTCCTTAAAAAGGGGAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAACAG 1064
Db 1077 -----AspSerCysPheAspProTyrThrValSerHis----- 1087
QY 1065 TTCACCATCAATACGTTGATGTCGATACCAACAATTTGCTAAAGAGTGAAGCTCTTA 1124
Db 1088 -----TyrAlaValGlyAspGluTyrGluArgMet---SerGluSerGlyPhe 1102
QY 1125 ACAGTAGCGAAGTAACTTAGACTTCAGAGATTTATACGATCCTCGTATAGGCTAAA 1184
Db 1103 LysLeuLeuCysGlnCysLeuGlyPheGlySerGlyHisPheArgCysAspSerSerArg 1122
QY 1185 CTACTCTACAACAATCTCGATGCTTTGGTATTATGACTATACCTTAACTGGAAAAAGTA 1244
Db 1123 TrpCysHisAspAsn-----GlyVal---AsnTyr-----LysIle 1133
QY 1245 GAGGATAATCAGCATGACACCAACCGGTATCATACCGTTTATATGGCAAGCGACCGAA 1304
Db 1134 GlyGluLysTyrAsp-----ArgGln 1140
QY 1305 GAGAGAAATGCT-----AGCTATCATTTAGCGGTGGTGTGTCAGCGCAGCAAAATG 1355
Db 1141 GlyGluAsnGlyGlnMetMetSerCysThrCysLeuGlyAsnGlyLysGlyGluPheLys 1160
QY 1356 GTTCAGCCCCAGTCCCGGTGGTGTGTCAGTCAAGCAAGCCCGGTGTTATGACAATGA 1415
Db 1161 CysAspProHisGluAlaThr-----CysTyrAspAspGly 1172
QY 1416 AAACACTATCAGATAAATCAACAGTGGGAGCGACCTTACCTAGTATATGTTGTTGT 1475
Db 1173 LysThrTyrHisValGlyGluGlnTyrGlnLysGluTyrLeuGlyAlaIleCysSerCys 1192
QY 1476 ACTTGTATTGAGGAAGCCGAGTTTAACTGCAAAAGT-----AAA 1517
Db 1193 ThrCysPheGlyGlyGlnArgGlyTyrArgCysAspAsnCysArgArgProGlyGlyGlu 1212
QY 1518 CTGAAGCTGAAGAGACTGCTTTTGACAAGTACACTGGGAACACTTAC 1565
Db 1213 ProSerProGluGlyThr-----ThrGlyGlnSerTyr 1223

```

RESULT 15

```

US-10-995-561-628
; Sequence 628, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702

```

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 628
; LENGTH: 1286
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-628

Alignment Scores:
Pred. No.: 9,09e-09 Length: 1286
Score: 211.50 Matches: 126
Percent Similarity: 33.28% Conservative: 72
Best Local Similarity: 21.18% Mismatches: 188
Query Match: 7.20% Indels: 209
DB: 6 Gaps: 27

US-09-940-235-10 (1-1661) x US-10-995-561-628 (1-1286)

QY 15 CTAGCCGGTCTCTCAACGACGAGGAGCAGATCATCGCGCCGCGGCGCAGGACCAAGC 74
Db 789 LeuIleGlyThrGlnSerThrAlaIleProAlaProThrAspLeuLysPheThrGlnVal 808
QY 75 TGCCCGAGATCTCGATCCCGGAA-----ATTAATACGACTCACTATAGG 119
Db 809 ThrProThrSerLeuSerAlaGlnTyrThrProProAsnValGlnLeuThrGlyTyrArg 828
QY 120 GAGACCAACACGGTTTCCCTCTAGAAAATAATTTGTTTAACTTTAAGAAGGAGATATACC 179
Db 829 ValArgValThrProLys----- 834
QY 180 ATGATGCTGACCTGAGTGGCTGTAGACGCTCATCTGTCAACAACAGCAATGTTGTT 239
Db 835 GluLysThrGlyProMetLysGluIleAsnLeuAlaProAspSerSerSer- 851
QY 240 GTTAGCCTTCTGCTACTGTTGAGGGGAGCAATCAAGACATTAAGTCTTAATTTTGTAA 299
Db 852 ValValValSerGlyLeuMetValAlaThrLysTyrGluValSerValTyrAlaLeuLys 871
QY 300 ATCGATCTAACATCAGCAGCTGCTCATGGAGGAAAGCAGAG---CAAGGCTTAAGTCCA 356
Db 872 AspThrLeuThrSerArgProAlaGlnGlyValValThrThrLeuGluAsnValSerPro 891
QY 357 AAATCAAAACCATTTGCTACTGATAGTGGCGCGATGTACATAAACTTGAGAAAGCTGAC 416
Db 892 ProArgAlaArgValThrAsp----- 899
QY 417 TTACTAAAGGCTATTCAAGAACAAATGATCGCTAAGTCCACAGTAAACGACGACTCTTT 476
Db 899 ----- 899
QY 477 GAGTCTATTGATTTTTCAGCGATGCAACCAATTAAGTATGATCGAAACGGCAAGGCTACTTT 536
Db 900 -----AlaThrGluThrIleThrIle----- 907
QY 537 GCTGACAAAGATGTTTCGTAACCTTCGCGACCAACCTGTCACCAAGATTTTGTGTAAGC 596
Db 908 -----SerTyrArgThrLysThrLysThrLysThrLysThrLysThrLysThrLys 922
QY 597 GGACATGTGCGGTAGACCATATATAAAGAAAAACCAATACAAAACCAAGCGAAATCTGTT 656
Db 923 Ala-----ValProAlaAsnGlyGlnThrProIleGlnArgThrIleLysProAsp 939
QY 657 GATGTGAATATACTGTACAGTTTACTCCCTTAAACCTGTATGAGATTTCAGACCAAGGT 716
Db 940 ValArgSerTyrThrIle-----ThrGlyLeuGlnProGlyThrAspTyrLys----- 955
QY 717 CTCAAAGATACTAAGTATTGAAACACACTAGCTATCGGTGACACCATCACATCTCAAGAA 776
Db 956 -----IleTyrLeuTyrThrLeuAsnAspAsnAlaArgSerSerPro 969
QY 777 TTAAGTCTCAAGCA----- 791
Db 970 ValValIleAspAlaSerThrAlaIleAspAlaProSerAsnLeuArgPheLeuAlaThr 989

```

```
QY 792 -----CAAAGCATTTTAAACAAAAACCCCA-----GGCTATACG 827
Db |||:|||||
990 ThrProAsnSerLeuLeuValSerTrpGlnProArgAlaArgIleThrGlyTyrIle 1009
QY 828 ATT---TATGACGTCACCTCCTCAATCGTCACATGACACATGACATTTTCCGTACGATT 884
Db |||:|||||:|
1010 IleLysTyrGluLysProGlySerProPro-----ArgGluVal 1022
QY 885 TTACCAATGGATCAA----- 899
Db |||:|||||
1023 ValProArgProArgProGlyValThrGluAlaThrIleThrGlyLeuGluProGlyThr 1042
QY 900 GAGTTTACTTACCGTGT-----AAAAATCGGGAACAAGCTTTATAGATC----- 944
Db |||:|||||
1043 GluTyrThrIleTyrValIleAlaLeuLysAsnAsnGlnLysSerGluProLeuIleGly 1062
QY 945 AATAAAAAATCTGGTCTGAATGAAGAAATAACAACACTGACCTGATCTCTGAGAAATAT 1004
Db |||:|||||:|
1063 ArgLysLysThrGly---GlnGluAlaLeuSerGlnThrIle----- 1076
QY 1005 TACGTCCTTAAAAAGGGGAAAGCGTATGATCCCTTTGATCGCAGTCACCTTGAACACTG 1064
Db |||:|||||
1077 -----SerTrpAlaProPheGln----- 1082
QY 1065 TTCACCATCAATACGTTGATCGATACCAACGAATTCGTAATAAAGTGAGCAGCTCTTA 1124
Db |||:|||||:|
1083 -----AspThrSerGluTyrIleIleSerCysHisProVal 1094
QY 1125 ACAGTAGCGAAGTAACCTTACAGCTTACAGATTTATACGATCCTCGTGATAAGGCTAAA 1184
Db |||:|||||
1095 GlyThrAspGluGluProLeuGlnPheArg-----ValProGlyThrSerThrSer 1111
QY 1185 CTACTCTACAACAATCTCGATCGCTTTGTGTATATGACTATACCTTAACCTGGAAGATA 1244
Db |||:|||||
1112 Ala-----ThrLeuThrGlyLeuThr 1118
QY 1245 GAGGATAATCACGATGACACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAA 1304
Db |||:|||||
1118 ----- 1118
QY 1305 GGAGAGAAATGTAGTATCAT---TTAGCCGGTGGTGGTCAGGCGCAGCAAAATGGTTCAG 1361
Db |||:|||||:|
1119 ---ArgGlyAlaThrTyrAsnIleIleValGluAlaLeuLysAspGlnGlnArgHisLys 1137
QY 1362 CCCGAGTCCCGGTGCTGTCTAGTCACCAAGCCCGGTGTGTATGACAAATGGAACAC 1421
Db |||:|||||
1138 ValArgGluGluValValThrValGlyAsnSerGlyTrpCysHisAspAsnGlyValAsn 1157
QY 1422 TATCAGATAAATCAACAGTGGGAGCGGACC---TACCTAGGTAAATGTGTGGTTGTACT 1478
Db |||:|||||:|
1158 TyrLysIleGlyGluLysTrpAspArgGlnGlyGluAsnGlyGlnMetMetSerCysThr 1177
QY 1479 TGTTATGAGGAGCCGAGGT---TTTAACTCGAAAGTAAACCTGAAGCTGAAGAGACT 1535
Db |||:|||||
1178 CysLeuGlyAsnGlyLysGlyGluPheLysCysAspProHis-----GluAlaThr 1194
QY 1536 TGCTTTGACAGTACACTGGGAACACTTACGAGTGGGTGACACACTTATGAGCGTCTCTAAA 1595
Db |||:|||||
1195 CysTyrAsp-----AspGlyLysThrTyrHisValGlyGluGlnTrpGlnLysGluTyr 1212
QY 1596 GACTCCATGATCGGAGCTGTACCTGATCGGGGCTGGCGGAGG 1640
Db |||:|||||
1213 LeuGlyAlaIleCysSerCysThrCysPheGlyGlyGlnArgGly 1227
```


GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 28, 2006, 01:49:35 ; Search time 25.4546 Seconds
(without alignments)
12556.953 Million cell updates/sec

Title: US-09-940-235-10
Perfect score: 2939
Sequence: 1: gcaacccgcagcctagcc.....gaataagctgtaccatctaa 1661

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p.model -DEV=xlp
-Q=/cgn2_1/USPTC_spool_p/US09940235/runat_27012006_144218_27578/app_query.fasta_1.7708
-DB=PIR -QFMT=fastan -SUFFIX=rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THRM SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09940235 @CGN 1 1 185 @runat_27012006_144218_27578 -NCPU=6 -ICPU=3
-NO_MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR 80.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1984	67.5	440	1 A22801	streptokinase prec
2	1951.5	66.4	415	1 B280	streptokinase (EC
3	1942	66.1	440	2 S02723	streptokinase G pr
4	1779	60.5	440	2 S02724	streptokinase A pr
5	1673	56.9	414	2 J00292	streptokinase - St
6	1670	56.8	440	2 S04168	streptokinase A pr
7	667	22.7	197	2 S53334	streptokinase - St
8	620	21.1	128	2 S77671	streptokinase A (E
9	600	20.4	2386	1 FNHU	fibronectin precu
10	575	19.6	2265	1 FNBO	fibronectin - bovl
11	574	19.5	2477	2 S14428	fibronectin precu
12	569	19.4	128	2 S77680	streptokinase A (E
13	557	19.0	128	2 S77688	streptokinase A (E
14	552	18.8	128	2 S77679	streptokinase A (E

15	496	16.9	2481	2 A43908	fibronectin - Afri
16	423	14.4	128	2 S77676	streptokinase A (E
17	419	14.3	128	2 S77673	streptokinase A (E
18	419	14.3	128	2 S77687	streptokinase A (E
19	417	14.2	128	2 S77682	streptokinase A (E
20	416	14.2	128	2 S77678	streptokinase A (E
21	408	13.9	128	2 S77677	streptokinase A (E
22	407	13.8	128	2 S77685	streptokinase A (E
23	401	13.6	128	2 S77683	streptokinase A (E
24	398	13.5	128	2 S77686	streptokinase A (E
25	397	13.5	128	2 S77684	streptokinase A (E
26	396	13.5	128	2 S77681	streptokinase A (E
27	390	13.3	128	2 S77674	streptokinase A (E
28	386	13.1	128	2 S77675	streptokinase A (E
29	171	5.8	1020	2 A29355	fibronectin - chic
30	164	5.6	103	2 A49173	fibronectin - mous
31	147.5	5.0	2231	2 D71870	hypothetical prote
32	144	4.9	190	2 J51279	fibronectin - east
33	133.5	4.5	823	2 A90761	hypothetical prote
34	132	4.5	732	2 B90975	hypothetical prote
35	130.5	4.4	1315	2 T28679	fibrinogen-binding
36	130.5	4.4	1745	2 A46431	tight junction-ass
37	129	4.4	823	2 F85624	hypothetical prote
38	128.5	4.4	860	2 T39502	hypothetical prote
39	125	4.3	503	2 JQ1019	site-specific DNA-
40	125	4.3	823	2 B85822	hypothetical prote
41	125	4.3	4688	2 F82885	hypothetical prote
42	124.5	4.2	967	2 S25081	phosphoenolpyruvat
43	124.5	4.2	1237	2 AC1583	internalin protein
44	124	4.2	2334	2 S32920	cell wall-associa
45	123.5	4.2	953	2 C89824	hypothetical prote

ALIGNMENTS

RESULT 1
A22801
streptokinase precursor - Streptococcus "equisimilis"
C;Species: Streptococcus "equisimilis"
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: A22801
R;Malke, H.; Roe, B.; Ferretti, J.J.
Gene 34, 357-362, 1995
A;Title: Nucleotide sequence of the streptokinase gene from Streptococcus equisimilis H
A;Reference number: A22801; MUID:85232082; PMID:2989113
A;Accession: A22801
A;Molecule type: DNA
A;Residues: 1-440 <MAL>
A;Cross-references: UNIPROT:P00779; UNIPARC:UPI000002BE73; GB:X72832; NID:g407876; PIDN:
A;Experimental source: strain H46A
C;Genetics:
A;Gene: skc
C;Superfamily: streptokinase

Alignment Scores:
Pred. No.: 3.95e-126 Length: 440
Score: 1984.00 Matches: 383
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 67.51% Indels: 0
DB: 1 Gaps: 0

US-09-940-235-10 (1-1661) x A22801 (1-440)

QY	183	ATTGCTTGACCTGAGTGGCTGCTAGACCGTCATCTGTCAACAACAGCAATTCGTTGTT	242
Db	27	ILaLaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal	46
QY	243	AGCGTTGCTGGTACTGTGTGAGGGGACGAATCAAGACATTAGTCTTAATATTTTGAATC	302
Db	47	SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle	66
QY	303	GATCTAACATCAGCACCTGCTCATGGAGGAAAGACAGACGACGAGGCTTAAGTCCAAATCA	362

```

Db      67 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 86
QY      363 AAACCAATTTGCTACTGATAGTCGGCGGATGTACATAAACTTCAGAAAGCTGACTACTA 422
Db      87 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 106
QY      423 AAGGCTATTCAAGAACAAATTGATCGCTAACGTCACAGTCCACAGTAACGACGACTACTTTGAGGTC 482
Db      107 LysAlaLleGlnGlnGlnLeuLeuAlaAsnValHisSerAsnAspAspTyrPheGluVal 126
QY      483 ATTGATTTTTCGAAGCGATCAACCACTACTGATCGAAACGGCAAGGCTCTACTTTGCTGAC 542
Db      127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146
QY      543 AAAGATGGTTGGTAACTTCGTCGGACCCCAACCTGTCCAAGAATTTTTCGCTAAACGGGACAT 602
Db      147 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 166
QY      603 GTGCGCGTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTGATGTCG 662
Db      167 ValArgValArgProTyrLysGluLysProLleGlnAsnGlnAlaLysSerValAspVal 186
QY      663 GAATATACCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTCAAA 722
Db      187 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 206
QY      723 GATACCTAGCTTATGAAACACTAGCTATCGGTGACACCATCACTCTCAAGAAATACTA 782
Db      207 AspThrLysLeuLysThrLeuAlaLleGlyAspThrIleThrSerGlnGluLeuLeu 226
QY      783 GCTCAAGCACAAAGCATTTTAAACAAAACCCAGGCTATACGATTTATGAACTGAC 842
Db      227 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 246
QY      843 TCCTCAATCGTCACATCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 902
Db      247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
QY      903 TTTACTACCGGTGTTAAATAACCGGAACAAAGCTTATAGGATCAATAAAAAATCTGCTG 962
Db      267 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 286
QY      963 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTTACGTCCTTAAAAAGGG 1022
Db      287 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 306
QY      1023 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACCTTCAAACTGTTCAACCATCAATACGTT 1082
Db      307 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326
QY      1083 GATGTCGATACCAACGAATTGCTAAAGATGAGCAGCTCTTAAACAGCTACGCAACGTAAC 1142
Db      327 AspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
QY      1143 TTAGACTTCAGAGATTTATACCATCTCTGTGATAGGCTTAAACTACTCTACACATCTC 1202
Db      347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
QY      1203 GATGCTTTTGGTATTATGACTATACCTTAACTTGAAGAAAGTAGAGGATATACGATGAC 1262
Db      367 AspAlaPheGlyIleLewcAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 386
QY      1263 ACCAACCGTATCATACCGTTTATATGGCAAGCGACCGAAGGAGAGAAATGCTAGCTAT 1322
Db      387 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 406
QY      1323 CATTTAGCC 1331
Db      407 HisLeuAla 409

```

RESULT 2
BZSO

```

streptokinase (EC 3.4.-.-) - Streptococcus sp.
C;Species: Streptococcus sp.
C;Date: 05-Apr-1983 #sequence_revision 05-Apr-1983 #text_change 09-Jul-2004
C;Accession: A00967
R;Jackson, K.W.; Tang, J.
Biochemistry 21, 6620-6625, 1982
A;Title: Complete amino acid sequence of streptokinase and its homology with serine protease
A;Reference number: A00967; MUID:83127125; PMID:6760891
A;Accession: A00967
A;Molecule type: protein
A;Residues: 1-415 <JAC>
A;Cross-references: UNIPROT:P00779; UNIPARC:UPI0000172BF5
A;Note: 169-Asp and 181-Asp were also found
A;Note: this protein is not a protease, but it activates plasminogen by complexing with
C;Superfamily: streptokinase
C;Keywords: hydrolase

Alignment Scores:
Pred. No.: 6,09e-124 Length: 415
Score: 1951.50 Matches: 380
Percent Similarity: 99.22% Conservative: 1
Best Local Similarity: 98.96% Mismatches: 2
Query Match: 66.40% Indels: 1
DB: 1 Gaps: 1

US-09-940-235-10 (1-1661) x BZSO (1-415)

```

```

QY      183 ATTGCTGGACTGAGTGGCTGTAGACCGTCCATCTCTCAACACACGCCAATTCGTTGTT 242
Db      1 IleAlaGlyProGluThrLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
QY      243 AGCGTTGCTGTTACTGTTGAGGGACGAATCAAGACATTAAGTCTTAATATTTTGAATC 302
Db      21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPheGluIle 40
QY      303 GATCTAACATCAACGACCTGCTCATGAGGAAAGACAGACGAAGCTTTAAGTCCAAAATCA 362
Db      41 AspLeuThrSerArgProAlaHisGlyLysThrGluGlnGlyLeuSerProLysSer 60
QY      363 AAACCAATTTGCTACTGATGTCGGCGATGTCATAAACTTGAGAAAGCTGACTACTA 422
Db      61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 80
QY      423 AAGCTATTCAAGACAACTGATCGCTAACGTCACAGTAACGACGACTACTTTGAGGTC 482
Db      81 LysAlaLleGlnGlnGlnLeuLeuAlaAsnValHisSerAsnAspAspTyrPheGluVal 100
QY      483 ATTGATTTTTCGAAGCGATGCAACCATTTACTGATCGAAACGGCAAGGCTCTACTTTGCTGAC 542
Db      101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
QY      543 AAAGATGGTTGCGTAAACCTTTGCGGACCCCAACCTGTCTCAAGAAATTTTTCGTAAGCGGACAT 602
Db      121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
QY      603 GTGCGGTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTCG 662
Db      141 ValArgValArgProTyrLysGluLysProLleGlnAsnGlnAlaLysSerValAspVal 160
QY      663 GAATATACCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTTCAGACGAGTCTCAAA 722
Db      161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
QY      723 GATACCTAGCTTATGAAACACTAGCTATCGGTGACACCATCACTCATCTCAAGAAATACTA 782
Db      181 LeuThrLysLeuLeuLysThrLeuAlaLleGlyAspThrIleThrSerGlnGluLeuLeu 200
QY      783 GCTCAAGCACAAAGCATTTTAAACAAAACCAAGCGGCTATACGATTTATGAAACGTCAC 842
Db      201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 220
QY      843 TCCTCAATCGTCACATCATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 902

```

Db	221	SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu	240
Qy	903	TTTACTACCGTGTAAAAATCGGAACAAAGCTTATAGAGTCAATAAAAAATCTGGTCTG	962
Db	241	PheThrTyArgValLysAsnArgGluGlnAlaTyArgIleAsnLysLysSerGlyLeu	260
Qy	963	AATGAAGAATAAACAACACTGACCTGATCTCTGAGAAATAT---TACGTCCTTAAAAA	1019
Db	261	AsnGluGluIleAsnAsnThrAspLeuIleSerLeuGluTyLysTyValLeuLysLys	280
Qy	1020	GGGGAAGACCGTATCATCCCTTTGATCGCAGTCACCTTGAACCTGTTCAACATCAATAC	1079
Db	281	GlyGluLysProTyAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTy	300
Qy	1080	GTTGATGTCATACCAACGAATTTGCTAAAAGTCAGCAGCTCTTAAACAGCTAGCGAACGT	1139
Db	301	ValAspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArg	320
Qy	1140	AACCTAGACTTCAGAGATTTATACGATCTCTGCGTGAATAAGCTAAACTTACTCTACAACAT	1199
Db	321	AsnLeuAspPheArgAspLeuTyAspProArgAspLysAlaLysLeuLeuTyAsnAsn	340
Qy	1200	CTCGATGCTTTGGTATTATGACTATACCTTAACCTGCGAAAGTAGAGGATATACGAT	1259
Db	341	LeuAspAlaPheGlyIleMetAspTyThrLeuThrGlyLysValGluAspAsnHisAsp	360
Qy	1260	GACACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAGGAGAGATGCTAGC	1319
Db	361	AspThrAsnArgIleIleThrValTyMetGlyLysArgProGluGlyGluAsnAlaSer	380
Qy	1320	TATCAATTTAGCC 1331	
Db	381	TyrHisLeuAla 384	
RESULT 3			
S02723			
streptokinase G precursor - Streptococcus sp. (strain 19908)			
C:Species: Streptococcus sp.			
C:Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 09-Jul-2004			
C:Accession: S02723			
R:Walter, F.; Siegel, M.; Maik, H.			
Nucleic Acids Res. 17, 1262, 1989			
A:Title: Nucleotide sequence of the streptokinase gene from a group-G Streptococcus.			
A:Reference number: S02723; MUID:89160265; PMID:2922269			
A:Accession: S02723			
A:Molecule type: DNA			
A:Residues: 1-440 <WAL>			
A:Cross-references: UNIPROT:P10519; UNIPARC:UPI000013614B; EMBL:X13400; NID:947095; PIDN			
C:Genetics:			
A:Gene: skg			
C:Superfamily: streptokinase			
F:1-26/Domain: signal sequence #status predicted <SIG>			
F:27-440/Product: streptokinase #status predicted <MAT>			
Alignment Scores:			
Pred. No.:	2,67e-123	Length:	440
Score:	1942.00	Matches:	375
Percent Similarity:	98.43%	Conservative:	2
Best Local Similarity:	97.91%	Mismatches:	6
Query Match:	66.08%	Indels:	0
DB:	2	Gaps:	0
US-09-940-235-10 (1-1661) x S02723 (1-440)			
Qy	183	ATTGCTGGACCTGAGTGGCTGCTAGACCGCTCCATCTGTCAACACAGCAATTTGGTTGT	242
Db	27	IleAlaGlyProGluTripleLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal	46
Qy	243	ACGCTGCTGGTACTGTTGAGGGGAGCAATCAAGACATTTAGTCTTAAATTTTGTGAATC	302
Db	47	SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle	66
Qy	303	GATCTAATCATCAGCACTCTCATGGAGGAAAGACAGACAGGCTTAAGTCCAAATCA	362

QY	Db	Seq1	Seq2	Score	Length	Matches	Conservative	Mismatches	Indels	Gaps
247	Db	SerSertIleValThrHisAspAsnAspIlePheargThrIleLeuProMetAspGlnGlu	266		414	414				
903	QY	TTTACTTACCGTGTAAATAATCGGGAACAAGCTTATAGCATCAATAAAAAAATCTGGTCTG	962		1673	324	23	36		
267	Db	PheThrTyrrHisVallysAsnArgGluGlnAlaTyrrGluIleAsnProLysThrGlyIle	286		90	60%	23	36		
963	QY	AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCTTAAAAAAGG	102		84	60%	23	36		
287	Db	LysGluLysThrAsnAsnThrAspLeuValSerGluLysTyrrTyrrValLeuLysGlnGly	306		84	60%	23	36		
1023	QY	GAAGAAGCGGTATGATCCCTTTGATCGCAGTCACATTGAAACTGTTCCACCATCAAAATCGTT	108		84	60%	23	36		
307	Db	GluLysProTyrrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrrVal	326		84	60%	23	36		
1083	QY	GATGTCGATACCAACGAATTGCTAAAAAGTCAGCAGCTCTTAAACAGCTAGCGCAACGTAAC	114		84	60%	23	36		
327	Db	AspValAsnThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluGlyAsn	346		84	60%	23	36		
1143	QY	TTAGACTTCAGAGATTTATACGATCCCTCGTGATAGGCTAAACTACTCTACAAACAATCTC	120		84	60%	23	36		
347	Db	LeuAspPheargAspLeuTyrrAspProArgAspLysAlaLysLeuLeuTyrrAsnAsnLeu	366		84	60%	23	36		
1203	QY	GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAGCTAGAGGATAATCACGATGAC	126		84	60%	23	36		
367	Db	AspAlaPheAspIleMetAsnTyrrThrLeuThrGlyLysValGluAspAsnHisAspLys	386		84	60%	23	36		
1263	QY	ACCAACCGTATCATACACCGTTTATATGGCAAGCGACCCGAGGAGAGAGAACTGCTAGCTAT	132		84	60%	23	36		
387	Db	AsnAsnArgValValThrValTyrrMetGlyLysArgProLysGlyAlaLysGlySerTyrr	406		84	60%	23	36		
1323	QY	CATTTAGGCC	1331		84	60%	23	36		
407	Db	HisLeuAla	409		84	60%	23	36		
RESULT 5										
JU0292										
C;Species:		Streptococcus pyogenes								
C;Date:		30-Jun-1992 #sequence_revision	30-Jun-1992	#text_change	09-Jul-2004					
C;Accession:		A43867	JU0292							
R;Okuni, H.; Todome, Y.; Suzuki, H.; Mizuse, M.; Kotani, N.; Horiuchi, K.; Sh										
A;Title:		Immunochemical studies and complete amino acid sequence of the strept								
A;Reference number:		A43867	MUID:	92104686	PMID:	1370275				
A;Accession:		A43867								
A;Molecule type:		protein								
A;Residues:		1-414	<ORF>							
A;Cross-references:		UNIPROT:Q57391; UNIPARC:UP10000175C82								
A;Experimental source:		M type 12 strain A374								
A;Note:		sequence extracted from NCBI backbone (NCBI:74592)								
C;Superfamily:		streptokinase								
Alignment Scores:										
Pred. No.:		3.55e-105								
Score:		1673.00								
Percent Similarity:		90.60%								
Best Local Similarity:		84.60%								
Query Match:		56.92%								
DB:		2								
US-09-940-235-10 (1-1661) x JU0292 (1-414)										
QY	183	ATTGCTGGACCTGATGGCTGCTAGACCGTCCATCTGTCAACACAGCCCAATTGGTTGT	242		414	414				
Db	1	IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal	20		414					

Db 41 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 60
 QY 363 AAACCACTTGTCTACTGATAGTGGCGCATGTCAATAAATCTGAGAAGCTGACTACTA 422
 Db 61 LysProPheAlaThrAsnSerSerAlaMetProHisLysLeuGluLysAlaAspLeu 80
 QY 423 AAGCTATTCAAGAACAAATTGATCGCTAAAGTCCACAGTAAACGACGACTACTTTGAGTGC 482
 Db 81 LysAlaIleGlnGluLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal 100
 QY 483 ATTGATTTCGACAGCGATGCAACCACTTACTGATCGAAGCGGACGCTTCTGCTGAC 542
 Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
 QY 543 AAAGATGGTTTCGGTAACTTCGCGACCAACCTGTCCAAAGAAATTTTGTCTAAGCGGACAT 602
 Db 121 ArgAspSerValThrLeuProThrGlnProValGlnGluPheLeuSerGlyHis 140
 QY 603 GTGCGGTTAGACATATAAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 662
 Db 141 ValArgValArgProTyrGlnProLysAlaValHisAsnSerAlaGluArgValAsnVal 160
 QY 663 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTCAGACCGAGTCTCAA 722
 Db 161 AsnTyrGluValSerPheValSerGluThrGlyAsnLeuAspPheThrProSerLeuLys 180
 QY 723 GATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCTCAAGATTACTA 782
 Db 181 GluArgTyrHisLeuThrThrLeuAlaValIleGlyAspSerLeuSerGlnGluLeuAla 200
 QY 783 GCTCAAGCAAAAGCAATTTTAAACAAAACCCAGGCTATACGATTTATGACAGTGCAC 842
 Db 201 AlaIleAlaGlnPheIleLeuSerLysGluHisProAspTyrIleIleThrLysArgAsp 220
 QY 843 TCTCTAATCGTCTCATGACATGACATTTTCGTTACGATTTTACCAATGGATCAAGAG 902
 Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
 QY 903 TTTACTACCGTGTAAATCGGAACAGCTTATAGGATCAATAAAATCTGGTCTG 962
 Db 241 PheThrTyrHisIleLysAspArgGluGlnAlaTyrLysAlaAsnSerLysThrGlyIle 260
 QY 963 AATGAAGAAATAAACCACTGATCTCTGAGAAATATTACGCTCTTAAAAAGGG 1022
 Db 261 ValGluLysThrAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280
 QY 1023 GAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAATCTGAACTCAACATACGTT 1082
 Db 281 GluGluProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
 QY 1083 GATGTCGATACCAACCAATGCTTAAAGTGACGAGCTCTTAACAGCTAGCAAGCTAAC 1142
 Db 301 AspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320
 QY 1143 TTAGACTTCAGAGATTATACGATCTCTGAGTAAAGGCTAACTCTACACAATCTC 1202
 Db 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
 QY 1203 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAAAGTAGAGGATAATACGATGAC 1262
 Db 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAsnAsp 360
 QY 1263 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCGGAGAGAGATGCTAGCTAT 1322
 Db 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 380
 QY 1323 CATTTAGCC 1331
 Db 381 HisLeuAla 383
 RESULT 6
 S04168
 streptokinase A precursor - Streptococcus pyogenes (strain NZ131)

C:Species: Streptococcus pyogenes
 C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 09-Jul-2004
 C:Accession: S04168
 R:Huang, T.T.; Malke, H.; Ferretti, J.J.
 Mol. Microbiol. 3, 197-205, 1989
 A:Title: The streptokinase gene of group A streptococci: cloning, expression in Escheri
 A:Reference number: S04168; MUID:89343623; PMID:2668686
 A:Accession: S04168
 A:Molecule type: DNA
 A:Residues: 1-440 <HUA>
 A:Cross-references: UNIPROT:Q57391; UNIPARC:UPI00000BD04A; EMBL:X51517; NID:947437; PID:
 C:Genetics:
 A:Gene: ska
 C:Superfamily: streptokinase
 F:1-26/Domain: signal sequence #status predicted <SIG>
 F:27-440/Product: streptokinase A #status predicted <MAT>
 Alignment Scores:
 Pred. NO.: 5.68e-105 Length: 440
 Score: 1670.00 Matches: 324
 Percent Similarity: 90.34% Conservative: 22
 Best Local Similarity: 84.60% Mismatches: 37
 Query Match: 56.82% Indels: 0
 DB: 2 Gaps: 0
 US-09-940-235-10 (1-1661) x S04168 (1-440)
 QY 183 ATTCTGACCTGAGTGGCTGCTAGACGCTCCATCTGTCAACACAGCAATGTTGTTT 242
 Db 27 IleAlaGlyTyrGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 46
 QY 243 AGCGTCTGCTACTGTTGAGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 302
 Db 47 SerValAlaGlyThrValGluGlyThrAsnGlnGluIleSerLeuLysPheGluIle 66
 QY 303 GATCTAATCATCAGCACTGCTCATGAGAGAAAGACAGACGAGGCTTAAGTCCAAAATCA 362
 Db 67 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 86
 QY 363 AAACATTTGCTACTGATGATGCGCGATGTCAATAAATCTGAGAAGCTGACTACTA 422
 Db 87 LysProPheAlaThrAspLysGlyAlaMetSerHisLysLeuGluLysAlaAspLeu 106
 QY 423 AAGCTATTCAAGAACAAATGATCGCTAAACGTCACAGTACGACGACTACTTTGAGTGC 482
 Db 107 LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal 126
 QY 483 ATTGATTTCGACAGCGATGCAACCATTTACTGATCGAAACGCAAGGCTTACTTTGCTGAC 542
 Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146
 QY 543 AAAGATGGTTTCGGTAACTTCGCGACCAACCTGTCCAAAGAAATTTTGTCTAAGCGGACAT 602
 Db 147 LysAspAspSerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 166
 QY 603 GTGCGGCTTAGACATATAAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 662
 Db 167 ValArgValLysProTyrGlnProLysAlaValHisAsnSerAlaGluArgValAsnVal 186
 QY 663 GAATATCTCTACAGTTTACTCCCTTAAACCCCTGATGACGATTCAGACGAGTCTCAA 722
 Db 187 AsnTyrGluValSerPheValSerGluThrGlyAspLeuAspPheThrProLeuLeu 206
 QY 723 GATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCTCAAGATTACTA 782
 Db 207 AsnGlnTyrHisLeuThrThrLeuAlaValGlyAspSerLeuSerGlnGluLeuAla 226
 QY 783 GCTCAAGCAAAAGCAATTTTAAACAAAACCAACCGGCTATACGATTTATGACAGTGC 842
 Db 227 AlaIleAlaGlnPheIleLeuSerLysLysHisProAspTyrIleIleThrLysArgAsp 246
 QY 843 TCCTCAATCGTCTCATGACAAATGACATTTTCGTACGATTTTACCAATGGATCAAGAG 902

Db 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
QY 903 TTTACTACCGTGTAAATAACCGGAACAAGCTTATAGGATCAATAAAAAATCTGCTCG 962
Db 267 PheThrTyHisIleLysAspArgGluGlnAlaTyLysAlaAsnSerLysThrGlyIle 286
QY 963 AATGAAGAAATAACAACACACTGACCTGATCTCTGGAATAATTATACGTCCTTAAAAAGGG 1022
Db 287 GluGluLysThrAsnAsnThrAspLeuIleSerGluLysTyTyValLeuLysLysGly 306
QY 1023 GAAAAGCGTATGATCCCTTTGATCGCAGTCACTGAAACTGTGTCAACATCAAAATACGTT 1082
Db 307 GluLysProTyAspProPheAspArgSerHisLeuLysLeuPheThrIleAsnTyVal 326
QY 1083 GATGTCGATACCAAGAAATGCTTAAAGTGAAGCGCTCTTAACAGCTAGCGCAACGTAAC 1142
Db 327 AspValAsnThrAsnLysLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
QY 1143 TTAGACTTCAGAGATTATACGATCCTCGTGATAAGGCTTAAACTACTCTACAACTATCTC 1202
Db 347 LeuAspPheArgAspLeuTyAspProArgAspLysAlaLysLeuLeuTyAsnAsnLeu 366
QY 1203 GATGCTTTTGTTATATGACTATACCTTAACTGGAAGAGTAGAGATAATCACGATGAC 1262
Db 367 AspAlaPheGlyIleMetAspTyThrLeuThrGlyLysValGluAspAsnHisAspAsp 386
QY 1263 ACCAACCGTATACACCGTTTATATGGCAAGCGACCGAGGAGAGAGATCGTAGCTAT 1322
Db 387 ThrAsnArgIleIleThrValTyMetGlyLysArgProGluGlyGluAsnAlaSerTy 406
QY 1323 CATTTAGCC 1331
Db 407 HisLeuAla 409

RESULT 7

S53334
streptokinase - Streptococcus sp.
C;Species: Streptococcus sp.
C;Date: 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change 09-Jul-2004
C;Accession: S53334
R;Shi, G.Y.; Chang, B.I.; Chen, S.M.; Wu, D.H.; Wu, H.L.
Biochem. J. 304, 235-241, 1994
A;Title: Function of streptokinase fragments in plasminogen activation.
A;Reference number: S53334; MUID:95091634; PMID:7998939
A;Accession: S53334
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-197 <SH1>
A;Cross-references: UNIPROT:Q7M115; UNIPARC:UPI000017AC2C

Alignment Scores:

Pred. No.:	2,09e-37	Length:	197
Score:	667.00	Matches:	167
Percent Similarity:	44.71%	Conservative:	2
Best Local Similarity:	44.18%	Mismatches:	1
Query Match:	22.69%	Indels:	208
DB:	2	Gaps:	9

US-09-940-235-10 (1-1661) x S53334 (1-197)

QY 183 ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCCAAATGGTTGT 242
Db 1 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsn----- 14
QY 243 AGCGTGTGCTGCTACTGTTGAGGGGCGAATCAAGACATAGTCTTAAATTTTGAATC 302
Db 14 ----- 14
QY 303 GATCTAACATCACGACCTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAGTCCAAATCA 362
Db 15 -----Ser 15
QY 363 AAACCATTTGCTACTGATAGTGGCGGATGTCACATAAACTTGAGAAAGCTGACTTACTA 422

Db 16 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 35
QY 423 AAGGCTATTCAAGAAACAATTGATCGCTAAGCTCCACAGTAAACGACGACTACTTTGAGTC 482
Db 36 LysAlaIleGlnTrpGlnIleLeu----- 43
QY 483 ATTGATTTTGCAGCGATGCAACCATTTACTGATCGAAACGGCAAGGCTCTACTTTGCTGAC 542
Db 44 -----AsnGlyLysValTyPheAlaAsp 51
QY 543 AAAGATGGTTTCGGTAACTTTGCCGACCCAACCTGTCCAAGAAATTTTGTCTAAGCGGACAT 602
Db 52 LysAspGlySerValThr----- 57
QY 603 GTGCGGTTAGACCATATATAAGAAAAACCAATACAAAAACCAAGCGAAATCTGTTGATGTG 662
Db 58 -----GluLysProIleGlnAsnGlnAlaLysSerValAspVal 70
QY 663 GAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCAGGTCTCAA 722
Db 71 Glu----- 71
QY 723 GATACTAAGCTATTGAAACACACTAGCTATCGGTGACACCATCATCATCTCAAGAATTACTA 782
Db 71 ----- 71
QY 783 GCTCAAGCAACAAGCAATTTTAAACAAACACCCAGGCTATACGATTTATGATGAACGTGAC 842
Db 72 -----AsnHisProGlyTyThrIleTyGluArgAsp 82
QY 843 TCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 902
Db 83 SerSerIleVal-----ThrIleLeuProMetAspGlnGlu 94
QY 903 TTTTACTTACCGTGTAAAAAATCGGGAACAAGCTTATAGGATCAATAAAAAATCTTGCTCG 962
Db 95 PheThrTy-----SerGlyLeu 100
QY 963 AATGAAGAAATAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1022
Db 101 AsnGluGluIleAsnAsnThrAspLeuIleSer----- 111
QY 1023 GAAAGCCGTATGATCCCTTTGATCGCAGCTCACTTGAAACTGTTCCACCATCAATAGCTT 1082
Db 112 -----TyVal 113
QY 1083 GATGTCGATACCAACCAATTGCTTAAAGTGAGCAGCTCTTAACAGCTAGCGAACGTAAAC 1142
Db 114 AspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 133
QY 1143 TTAGACTTCAGAGATTATACGATCCTCGTGATTAAGGCTTAAACTACTCTACAACTATCTC 1202
Db 134 -----AspLeuTyAspProArgAspLysAlaLysLeuLeuTyAsnAsnLeu 149
QY 1203 GATGCTTTGTTGTTATGACTATACCTTAACTGGAAGAGTAGAGGATATACGATGAC 1262
Db 150 AspAlaPheGlyIleMet----- 155
QY 1263 ACCAACCGTATCATACCGTTTATATGGCAAGCGACCGCAAGGAGAGAAATGCT 1316
Db 156 -----IleIleThrValTyMetGlyLysArgProGluGlyGluAsnAla 170

RESULT 8

S77671
streptokinase A (EC 3.4.-.-) (alleles 2 and 3) - Streptococcus pyogenes (fragment)
C;Species: Streptococcus pyogenes
C;Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 22-Jun-1999
C;Accession: S77671; S77672
R;Kapoor, V.; Kanjilal, S.; Hamrick, M.R.; Li, L.L.; Whittam, T.S.; Sawyer, S.A.; Musser, M.L. Microbiol. 16, 509-519, 1995
A;Title: Molecular population genetic analysis of the streptokinase gene of Streptococcus
A;Reference number: S77671; MUID:96037795; PMID:7565111

A:Accession: S77671
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-128 <KAP>
A:Cross-references: UNIPARC:UPI00000000562; EMBL:U25853; NID:g818908; PIDN:AAA85729.1; PT:PT
A:Experimental source: strain ET1/M1
A>Note: allele 2
A:Accession: S77672
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-128 <KA2>
A:Cross-references: UNIPARC:UPI00000000562; EMBL:U25854; NID:g818910; PIDN:AAA85730.1; PT:PT
A:Experimental source: strain E2/M3
A>Note: allele 3
C:Genetics:
A:Gene: ska
C:Superfamily: streptokinase
C:Keywords: hydrolase; plasminogen activator; virulence

Alignment Scores:
Pred. No.: 3.01e-34 Length: 128
Score: 620.00 Matches: 118
Percent Similarity: 95.31% Conservative: 4
Best Local Similarity: 92.19% Mismatches: 6
Query Match: 21.10% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-10 (1-1661) x S77671 (1-128)

Qy	603	GTGCGGTTAGACCATATAAGAAAACCAATACAAAACGCGAAATCTGTTGATGTG	662
Db	1	ValArgValArgProTyrLysGluLysProValGlnAsnGlnAlaLysSerValAspVal	20
Qy	663	GAATATACTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTCAGACCAGGTCTCTCAA	722
Db	21	GlutThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys	40
Qy	723	GATACTAAGCTATTGAAAAACACTAGCTATCGTGACACCATCACATCTCAAGAATTACTTA	782
Db	41	AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu	60
Qy	783	GCTCAAGCAACAAGCATTTTAAACAAAACACCCAGGCTATACGATTTATGAACGTGAC	842
Db	61	AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp	80
Qy	843	TCCTCAATCGTCACTCATGACATGACATTTCCGTACGATTTTACCAATCGATCAAGAG	902
Db	81	SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu	100
Qy	903	TTTACTTACCGTGTTAAAAATCGGAAACAAGCTTATAGGACCAATAAAAAATCTGCTCTG	962
Db	101	PheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsnProLysThrGlyIle	120
Qy	963	AATGAAGAATAAACAACACTGAC	986
Db	121	LysGluLysThrAsnAsnThrAsp	128

RESULT 9
FNHU
fibronectin precursor [validated] - human
N:Alternate names: fibronectin splice form ED-A
C:Species: Homo sapiens (man)
C:Date: 27-Nov-1985 #sequence revision 31-Mar-1993 #text change 31-Dec-2004
C:Accession: A26460; A26284; S03917; A24854; A24476; A91008; A93529; A21011; A90495; A22
R:Dean, D.C.; Bowlin, C.L.; Bourgeois, S.
Proc. Natl. Acad. Sci. U.S.A. 84, 1876-1880, 1987
A:Title: Cloning and analysis of the promoter region of the human fibronectin gene.
A:Reference number: A26460; MUID:87175578; PMID:3031656
A:Accession: A26460
A:Molecule type: DNA
A:Residues: 1-49 <DEA>
A:Cross-references: UNIPROT:P02751; UNIPROT.Q14327; UNIPARC:UPI000016A926; GB:M15801; NT
R:Oldberg, A.; Ruoslahti, E.

A;Residues: 1948-2067 <UMB>
A;Cross-references: UNIPARC:UPI000000046A; GB:M27589; NID:g182705; PIDN:AAAS2465.1; PID:
A;Accession: B22245
A;Molecule type: mRNA
A;Residues: 1975-1991,2017-2039 <UM2>
A;Cross-references: UNIPARC:UPI0000174332; UNIPARC:UPI0000174333; GB:M27590
R;Sekiguchi, K.; Klos, A.M.; Kurachi, K.; Yoshitake, S.; Hakomori, S.
Biochemistry 25, 4936-4941, 1986
A;Title: Human liver fibronectin complementary DNAs: identification of two different mes
A;Reference number: I52394; MUID:87026578; PMID:3021206
A;Accession: I52723
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1978-1990,2016-2018,'N',2113-2127 <SEK>
A;Cross-references: UNIPARC:UPI00006E04C; GB:M14060; NID:g182701; PIDN:AAAS2464.1; PID:
R;Kornblitt, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
Proc. Natl. Acad. Sci. U.S.A. 80, 3218-3222, 1983
A;Title: Isolation and characterization of cDNA clones for human and bovine fibronectin
A;Reference number: A21165; MUID:83221567; PMID:6304699
A;Accession: A21165
A;Molecule type: mRNA
A;Residues: 2291-2386 <K03>
A;Cross-references: UNIPARC:UPI0000174334; GB:K00799; NID:g182681; PIDN:AAAS2460.1; PID:
R;Garcia-Pardo, A.; Pearlstein, E.; Frangione, B.
J. Biol. Chem. 258, 12670-12674, 1983
A;Title: Primary structure of human plasma fibronectin.
A;Reference number: A92398; MUID:84032463; PMID:6630202
A;Accession: A92398
A;Molecule type: protein
A;Residues: 32-47,'C',49-51,'S',53-72,'A',74-290 <GAR1>
A;Cross-references: UNIPARC:UPI0000174335
R;Garcia-Pardo, A.; Gold, L.I.
Arch. Biochem. Biophys. 304, 181-188, 1993
A;Title: Further characterization of the binding of fibronectin to gelatin reveals the R
A;Reference number: S34791; MUID:93312001; PMID:8323285
A;Accession: S34791
A;Molecule type: protein
A;Residues: 291-300:551-560 <GAR2>
A;Cross-references: UNIPARC:UPI0000174336; UNIPARC:UPI0000174337
R;Griffin, C.A.; Calaycay, J.; Shively, J.E.; Smith, R.L.
Thromb. Res. 43, 469-477, 1986
A;Title: Two plasma fibronectin fragments with different gelatin-binding properties.
A;Reference number: A60904; MUID:87019725; PMID:3532418
A;Accession: A60904
A;Molecule type: protein
A;Residues: 293-301 <GRI>
A;Cross-references: UNIPARC:UPI0000174338
R;Calaycay, J.; Pande, H.; Lee, T.; Borsi, L.; Siri, A.; Shively, J.E.; Zardi, L.
J. Biol. Chem. 260, 12136-12141, 1985
A;Title: Primary structure of a DNA- and heparin-binding domain (domain III) in human pl
A;Reference number: A23901; MUID:86008277; PMID:3900070
A;Accession: A23901
A;Molecule type: protein
A;Residues: 616-677,'Q',679-703,'PT' <CAL>
A;Cross-references: UNIPARC:UPI0000174339
R;Pierschbacher, M.D.; Ruoslahti, E.; Sundelin, J.; Lind, P.; Peterson, P.A.
J. Biol. Chem. 257, 9593-9597, 1982
A;Title: The cell attachment domain of fibronectin. Determination of the primary structu
A;Reference number: A92386; MUID:82265604; PMID:7050098
A;Accession: A92386
A;Molecule type: protein
A;Residues: 1441-1548 <PIE>
A;Cross-references: UNIPARC:UPI0000141CD5
A;Note: residues 1524-1527 are responsible for the cell-binding activity
R;Garcia-Pardo, A.; Rostagno, A.; Frangione, B.
Biochem. J. 241, 923-928, 1987
A;Title: Primary structure of human plasma fibronectin. Characterization of a 38 kDa dom
A;Reference number: A32517; MUID:87241275; PMID:3593230
A;Accession: A32517
A;Molecule type: protein
A;Residues: 1589-1630,'T',1722-2058 <GAR3>
A;Cross-references: UNIPARC:UPI000017433A
R;Tressel, T.; McCarthy, J.B.; Calaycay, J.; Lee, T.D.; Legesse, K.; Shively, J.E.; Pand

Biochem. J. 274, 731-738, 1991
A;Title: Human plasma fibronectin. Demonstration of structural differences between the A
A;Reference number: S14357; MUID:91190085; PMID:2012601
A;Accession: S14357
A;Molecule type: protein
A;Residues: 1614-1630,'T',1722-2081,2113-2244 <TRE>
A;Cross-references: UNIPARC:UPI000017433B
R;Garcia-Pardo, A.; Pearlstein, E.; Frangione, B.
J. Biol. Chem. 260, 10320-10325, 1985
A;Title: Primary structure of human plasma fibronectin. Characterization of a 31,000-da
A;Reference number: A23891; MUID:85261459; PMID:4019516
A;Accession: A23891
A;Molecule type: protein
A;Residues: 2071-2080;2112-2356 <GAR4>
A;Cross-references: UNIPARC:UPI000017433C; UNIPARC:UPI000017433D
C;Comment: The extra domain and connecting strand 3 are subject to developmental and tis
C;Comment: The cellular and plasma fibronectins are high molecular weight glycoproteins
ation, and transformation.
C;Genetics:
A;Gene: GDB:FN1
A;Cross-references: GDB:119135; OMIM:135600
A;Map position: 2q34-2q34
A;Introns: 49/3; 1266/1; 1357/1; 1447/1; 1487/1; 1541/1; 1631/1; 1721/1; 1991/1; 2145/1
C;Superfamily: fibronectin type I repeat homology; fibronectin type II repeat homology;
C;Keywords: acute phase; alternative splicing; cell adhesion; collagen binding; duplica
F;1-26/Domain: signal sequence #status predicted <SIG>
F;27-31/Domain: propeptide #status predicted <PRO>
F;32-2386/Product: fibronectin #status experimental <MAT>
F;52-272/Domain: fibrin and heparin binding <FHB>
F;52-87/Domain: fibronectin type I repeat homology <1F1>
F;97-135/Domain: fibronectin type I repeat homology <1F2>
F;141-179/Domain: fibronectin type I repeat homology <1F3>
F;186-225/Domain: fibronectin type I repeat homology <1F4>
F;231-270/Domain: fibronectin type I repeat homology <1F5>
F;308-608/Domain: collagen binding <CBR>
F;308-342/Domain: fibronectin type I repeat homology <1F6>
F;360-401/Domain: fibronectin type II repeat homology <2F1>
F;420-461/Domain: fibronectin type II repeat homology <2F2>
F;470-508/Domain: fibronectin type I repeat homology <1F7>
F;518-555/Domain: fibronectin type I repeat homology <1F8>
F;561-599/Domain: fibronectin type I repeat homology <1F9>
Alignment Scores:
Pred. No.: 7-72e-33 Length: 2386
Score: 600.00 Matches: 106
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.42% Indels: 0
DB: 1 Gaps: 0
US-09-940-235-10 (1-1661) x FNHU (1-2386)
QY 1341 CAGGCGCAGCAATGGTTTCAGCCCGTCCCGTGGCTCTCAGTCAAGCAAGCCCGGT 1400
Db 32 GlnAlaGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 51
QY 1401 TGTATGACATCGAAAAACACTATCAGATAAATCAACAGTGGGCGGACCTACTAGGT 1460
Db 52 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 71
QY 1461 AATGTGTGTGGTTGTTACTTGTATGAGGAAGCCGAGGTTTAACTCGGAAAGTAAACCT 1520
Db 72 AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnGlyGluSerLysPro 91
QY 1521 GAAGCTGAAGAGACTTGGCTTTTGACAAATACACTGGGAACACTTACCAGAGTGGGTGACACT 1580
Db 92 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 111
QY 1581 TATGAGCGTCTCTAAAGACTCATGATCTGGGACTGTGACCTGCGATCGGGGCTGGCGGAGG 1640
Db 112 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 131
QY 1641 AGAATAAGCTGTACCATC 1658

Db 132 ArgileSerCysThrille 137
RESULT 10
FNBO
F:2085-2124/Domain: fibronectin type I repeat homology <1F10>
F:2130-2167/Domain: fibronectin type I repeat homology <1F11>
F:2174-2209/Domain: fibronectin type I repeat homology <1F12>
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:3/Cross-link: isopeptide (Gln) (interchain to Lys N6-amino of fibrin) #status experimental
F:21-47,45-56,66-94,92-104,110-138,136-148,155-184,182-194,200-229,227-239,277-304,302-
7,2155-2167,2174-2200,2198-2209/Disulfide bonds: #status predicted
F:399,437,511,846,976,1213,1987/Binding site: carboxylate (Asn) (covalent) #status absent
F:1205,1692/Binding site: carboxylate (Asn) (covalent) #status absent
F:1943,1944/Binding site: carboxylate (Thr) (covalent) #status experimental
F:2246/Disulfide bonds: interchain (to 2250) #status predicted
F:2250/Disulfide bonds: interchain (to 2246) #status predicted
F:2263/Binding site: phosphate (Ser) (covalent) #status experimental
Alignment Scores:
Pred. No.: 3,73e-31 Length: 2265
Score: 575.00 Matches: 100
Percent Similarity: 97.17% Conservative: 3
Best Local Similarity: 94.34% Mismatches: 3
Query Match: 19.56% Indels: 0
DB: 1 Gaps: 0
US-09-940-235-10 (1-1661) x FNBO (1-2265)
Qy 1341 CAGCGCAGCAATGGTTTCAGCCCGGTCGCTCAGTCAAGCAAGCCCGGT 1400
Db 1 GlnAlaGlnGlnIleValGlnProGlnSerProLeuThrValSerGlnSerIleProGly 20
Qy 1401 TGTATGACAAATGGAAACACTATCAGATAAATCAACAGTGGGAGCGGACTTACTAGGT 1460
Db 21 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 40
Qy 1461 AATGTGTGGTTCCTACTTGTATGAGGAGCGGAGGCTTTTACTGCGAAGTAAACCT 1520
Db 41 SerAlaLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGluSerLysPro 60
Qy 1521 GNACTCAAGACACTGCTTTGACAAAGTACACTGGGACACTTACCGAGTGGGTGACACT 1580
Db 61 GluProGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
Qy 1581 TATGAGCGTCTTAAAGACTCATCATCTGGGACTGTACTGCATCGGGGCTGGCGGAGGG 1640
Db 81 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 100
Qy 1641 AGAATAAGCTGTACATC 1658
Db 101 ArgIleSerCysThrille 106
RESULT 11
S14428
fibronectin precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 28-Oct-1994 #sequence revision 28-Oct-1994 #text change 09-Jul-2004
C:Accession: S14428; S12455; A22319; S46203; S00459; A27252; I59049
R:Hynes, R.O.
submitted to the EMBL Data Library, July 1989
A:Reference number: S14428
A:Accession: S14428
A:Molecule type: mRNA
A:Residues: 1-2477 <HYN>
A:Cross-references: UNIPROT:P04937; UNIPARC:UPI000012A7C6; EMBL:X15906; NID:g56163; PID:
R:Schwarzbauer, J.E.; Patel, R.S.; Fonda, D.; Hynes, R.O.
EMBO J. 6, 2573-2580, 1987
A:Title: Multiple sites of alternative splicing of the rat fibronectin gene transcript.
A:Reference number: S12455; MUID:98054951; PMID:2445560
A:Accession: S12455
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 609-1810, 'T', 1812-2283 <SCH>
A:Cross-references: UNIPARC:UPI0000177AEA; EMBL:X15906
R:Tamkun, J.W.; Schwarzbauer, J.E.; Hynes, R.O.
Proc. Natl. Acad. Sci. U.S.A. 81, 5140-5144, 1984
A:Title: A single rat fibronectin gene generates three different mRNAs by alternative splicing

Db 132 ArgileSerCysThrille 137
RESULT 10
FNBO
F:2085-2124/Domain: fibronectin type I repeat homology <1F10>
F:2130-2167/Domain: fibronectin type I repeat homology <1F11>
F:2174-2209/Domain: fibronectin type I repeat homology <1F12>
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:3/Cross-link: isopeptide (Gln) (interchain to Lys N6-amino of fibrin) #status experimental
F:21-47,45-56,66-94,92-104,110-138,136-148,155-184,182-194,200-229,227-239,277-304,302-
7,2155-2167,2174-2200,2198-2209/Disulfide bonds: #status predicted
F:399,437,511,846,976,1213,1987/Binding site: carboxylate (Asn) (covalent) #status absent
F:1205,1692/Binding site: carboxylate (Asn) (covalent) #status absent
F:1943,1944/Binding site: carboxylate (Thr) (covalent) #status experimental
F:2246/Disulfide bonds: interchain (to 2250) #status predicted
F:2250/Disulfide bonds: interchain (to 2246) #status predicted
F:2263/Binding site: phosphate (Ser) (covalent) #status experimental
Alignment Scores:
Pred. No.: 3,73e-31 Length: 2265
Score: 575.00 Matches: 100
Percent Similarity: 97.17% Conservative: 3
Best Local Similarity: 94.34% Mismatches: 3
Query Match: 19.56% Indels: 0
DB: 1 Gaps: 0
US-09-940-235-10 (1-1661) x FNBO (1-2265)
Qy 1341 CAGCGCAGCAATGGTTTCAGCCCGGTCGCTCAGTCAAGCAAGCCCGGT 1400
Db 1 GlnAlaGlnGlnIleValGlnProGlnSerProLeuThrValSerGlnSerIleProGly 20
Qy 1401 TGTATGACAAATGGAAACACTATCAGATAAATCAACAGTGGGAGCGGACTTACTAGGT 1460
Db 21 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 40
Qy 1461 AATGTGTGGTTCCTACTTGTATGAGGAGCGGAGGCTTTTACTGCGAAGTAAACCT 1520
Db 41 SerAlaLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGluSerLysPro 60
Qy 1521 GNACTCAAGACACTGCTTTGACAAAGTACACTGGGACACTTACCGAGTGGGTGACACT 1580
Db 61 GluProGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
Qy 1581 TATGAGCGTCTTAAAGACTCATCATCTGGGACTGTACTGCATCGGGGCTGGCGGAGGG 1640
Db 81 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 100
Qy 1641 AGAATAAGCTGTACATC 1658
Db 101 ArgIleSerCysThrille 106
RESULT 11
S14428
fibronectin precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 28-Oct-1994 #sequence revision 28-Oct-1994 #text change 09-Jul-2004
C:Accession: S14428; S12455; A22319; S46203; S00459; A27252; I59049
R:Hynes, R.O.
submitted to the EMBL Data Library, July 1989
A:Reference number: S14428
A:Accession: S14428
A:Molecule type: mRNA
A:Residues: 1-2477 <HYN>
A:Cross-references: UNIPROT:P04937; UNIPARC:UPI000012A7C6; EMBL:X15906; NID:g56163; PID:
R:Schwarzbauer, J.E.; Patel, R.S.; Fonda, D.; Hynes, R.O.
EMBO J. 6, 2573-2580, 1987
A:Title: Multiple sites of alternative splicing of the rat fibronectin gene transcript.
A:Reference number: S12455; MUID:98054951; PMID:2445560
A:Accession: S12455
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 609-1810, 'T', 1812-2283 <SCH>
A:Cross-references: UNIPARC:UPI0000177AEA; EMBL:X15906
R:Tamkun, J.W.; Schwarzbauer, J.E.; Hynes, R.O.
Proc. Natl. Acad. Sci. U.S.A. 81, 5140-5144, 1984
A:Title: A single rat fibronectin gene generates three different mRNAs by alternative splicing

A;Reference number: A22319; MUID:84298097; PMID:6089177
 A;Accession: A22319
 A;Molecule type: DNA
 A;Residues: 2052-2237 <TAM>
 A;Cross-references: UNIPARC:UPI0000177ARB
 R;Falkenberg, C.; Enghild, J.J.; Thøgersen, I.B.; Salvesen, G.; Akerstroem, B. Biochem. J. 301, 745-751, 1994
 A;Title: Isolation and characterization of fibronectin-alpha(1)-microglobulin complex in A;Reference number: S46203; MUID:94330948; PMID:7519849
 A;Accession: S46203
 A;Status: preliminary
 A;Molecule type: protein
 A;Residues: 1183-1192; 'GLN', 1268, 'P', 1270-1271, 'D', 1273, 'CF', 1276, 'PY', 1385-1399 <PAL>
 A;Cross-references: UNIPARC:UPI0000177AEC; UNIPARC:UPI0000177AED; UNIPARC:UPI0000177AEE R;Patel, R.S.; Odermatt, E.; Schwarzbauer, J.E.; Hynes, R.O. EMBO J. 6, 2565-2572, 1987
 A;Title: Organization of the fibronectin gene provides evidence for exon shuffling during A;Reference number: S00459; MUID:88054950; PMID:3119323
 A;Accession: S00459
 A;Molecule type: DNA
 A;Residues: 1-139;2382-2477 <PAT>
 A;Cross-references: UNIPARC:UPI000017095C; UNIPARC:UPI0000177ABF; EMBL:X05831
 A;Note: the authors translated the codon CCT for residues 51 and 94 as Ala R;Schwarzbauer, J.E.; Tamkun, J.W.; Lemischka, I.R.; Hynes, R.O. Cell 35, 421-431, 1983
 A;Title: Three different fibronectin mRNAs arise by alternative splicing within the coding A;Reference number: A27252; MUID:84082067; PMID:6317187
 A;Accession: A27252
 A;Molecule type: mRNA
 A;Residues: 1586-1720, 'T', 1722, 1813-2477 <SC2>
 A;Cross-references: UNIPARC:UPI0000177AF0
 R;Odermatt, E.; Tamkun, J.W.; Hynes, R.O. Proc. Natl. Acad. Sci. U.S.A. 82, 6571-6575, 1985
 A;Title: Repeating modular structure of the fibronectin gene: Relationship to protein sub A;Reference number: 159049; MUID:86016741; PMID:3863113
 A;Accession: 159049
 A;Status: translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1722-1810 <RRS>
 A;Cross-references: UNIPARC:UPI000000040C; GB:M11750; NID:g204164; PIDN:AAA41170.1; PID: A;Genetics:
 C;Introns: 51/1; 94/1; 2416/3; 2454/3
 C;Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II repeat homology
 C;Keywords: alternative splicing; cell adhesion; collagen binding; disulfide bond; duplication
 F;1-32/Domain: signal sequence #status predicted <SIG>
 F;33-2477/Product: fibronectin #status predicted <MAT>
 F;53-88/Domain: fibronectin type I repeat homology <I>
 F;98-136/Domain: fibronectin type I repeat homology <I>
 F;142-180/Domain: fibronectin type I repeat homology <I>
 F;187-226/Domain: fibronectin type I repeat homology <I>
 F;232-271/Domain: fibronectin type I repeat homology <I>
 F;308-342/Domain: fibronectin type I repeat homology <I>
 F;360-401/Domain: fibronectin type II repeat homology <I>
 F;420-461/Domain: fibronectin type II repeat homology <I>
 F;470-508/Domain: fibronectin type I repeat homology <I>
 F;518-555/Domain: fibronectin type I repeat homology <I>
 F;561-599/Domain: fibronectin type I repeat homology <I>
 F;609-692/Domain: fibronectin type III repeat homology <I>
 F;718-800/Domain: fibronectin type III repeat homology <I>
 F;809-890/Domain: fibronectin type III repeat homology <I>
 F;905-987/Domain: fibronectin type III repeat homology <I>
 F;995-1076/Domain: fibronectin type III repeat homology <I>
 F;1085-1164/Domain: fibronectin type III repeat homology <I>
 F;1172-1257/Domain: fibronectin type III repeat homology <I>
 F;1265-1348/Domain: fibronectin type III repeat homology <I>
 F;1356-1439/Domain: fibronectin type III repeat homology <I>
 F;1447-1529/Domain: fibronectin type III repeat homology <I>
 F;1537-1619/Domain: fibronectin type III repeat homology <I>
 F;1614-1616/Domain: cell attachment (R-G-D) motif
 F;1631-1713/Domain: fibronectin type III repeat homology <I>
 F;1721-1803/Domain: fibronectin type III repeat homology <I>
 F;1811-1893/Domain: fibronectin type III repeat homology <I>
 F;1903-1984/Domain: fibronectin type III repeat homology <I>

F;1392-2074/Domain: fibronectin type III repeat homology <I>
 F;2181-2183/Region: cell attachment (R-G-D) motif
 F;2193-2273/Domain: fibronectin type III repeat homology <I>
 F;2296-2335/Domain: fibronectin type I repeat homology <I>
 F;2341-2378/Domain: fibronectin type I repeat homology <I>
 F;2385-2420/Domain: fibronectin type I repeat homology <I>
 F;53-79, 77-88, 98-126, 142-136, 142-170, 168-180, 187-216, 214-226, 232-261, 259-271, 308-335, 333-368, 2366-2378, 2385-2411, 2409-2420/Disulfide bonds: #status predicted
 F;2458/Disulfide bonds: interchain (to 2462) #status predicted
 F;2462/Disulfide bonds: interchain (to 2458) #status predicted
 Alignment Scores:
 Pred. No.: 4.37e-31 Length: 2477
 Score: 574.00 Matches: 100
 Percent Similarity: 97.17% Conservatives: 3
 Best Local Similarity: 94.34% Mismatches: 3
 Query Match: 19.53% Indels: 0
 DB: 2 Gaps: 0
 US-09-940-235-10 (1-1661) x S14428 (1-2477)
 QY 1341 CAGCGCAGCAATGTTTCAGCCCGTCCCGTGGCTGTCAGTCAAGCAAGCCCGT 1400
 Db 33 GlnAlaGlnGlnIleValGlnProSerProValAlaValSerGlnSerLysProGly 52
 QY 1401 TGTATGACATCGAAGAACATATCATCAATCAACAGTGGGACCGACCTAGGT 1460
 Db 53 CysPheAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 72
 QY 1461 AATGTGTGTTGTTGTTACTTGTATGAGGAGCCGAGGTTTAACTGCGAAAGTAAACCT 1520
 Db 73 AsnAlaLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGluSerLysPro 92
 QY 1521 GAAGCTGAAGAGACTGCTTTTGACAGTACACTGCGAACACTTACCGAGTGGTGACACT 1580
 Db 93 GluProGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrLysValGlyAspThr 112
 QY 1581 TATGAGGCTCTAAAGACTCATGATCTGGGACTGTACCTGCATCGGGGCTGGGCGAGG 1640
 Db 113 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 132
 QY 1641 AGAATAAGCTGTACCATC 1658
 Db 133 ArgLysCysThrIle 138
 RESULT 12
 S77680
 streptokinase A (EC 3.4.-.-) (allele 11) - Streptococcus pyogenes (fragment)
 C;Species: Streptococcus pyogenes
 C;Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004
 C;Accession: S77680
 R;Kapur, V.; Kanjilal, S.; Hamrick, M.R.; Li, L.L.; Whittam, T.S.; Sawyer, S.A.; Musser M.L. Microbiol. 16, 509-519, 1995
 A;Title: Molecular population genetic analysis of the streptokinase gene of Streptococcus A;Reference number: S77671; MUID:96037795; PMID:7565111
 A;Accession: S77680
 A;Status: translation not shown
 A;Molecule type: DNA
 A;Residues: 1-128 <KAP>
 A;Cross-references: UNIPROT:Q54687; UNIPARC:UPI00000BD2D3; EMBL:U25862; NID:g818926; PII: A;Experimental source: strain ET51/ML7
 C;Genetics:
 A;Gene: ska
 C;Superfamily: streptokinase
 C;Keywords: hydrolase; plasminogen activator; virulence
 Alignment Scores:
 Pred. No.: 8.23e-31 Length: 128
 Score: 569.00 Matches: 110
 Percent Similarity: 91.41% Conservatives: 7
 Best Local Similarity: 85.94% Mismatches: 11
 Query Match: 19.36% Indels: 0
 DB: 2 Gaps: 0

US-09-940-235-10 (1-1661) x S77680 (1-128)

```

Qy 603 GTGCGGTTAGACCATATAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 662
Db 1 ValArgValArgProTyrLysGluLysProIleGlnThrProAlaLysSerValAspVal 20
Qy 663 GAATATACCTGACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCAAGTCTCAAA 722
Db 21 ArgTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProValLeuLys 40
Qy 723 GATACCTAAGCTATTGAAACCACTAGCTATCGGTGACACCATCATCATCTCAAGAATTACTA 782
Db 41 AsnThrLysLeuLysThrLeuAlaIleGlyThrValThrSerGlnGluLeu 60
Qy 783 GCTCAAGCACAAGCATTTTAAACAAAACCAACCCAGCTATACGATTATGAACGTGAC 842
Db 61 AlaGlnAlaGlnSerIleLeuAsnGluSerHisProAspTyrThrIleTyrGluArgAsp 80
Qy 843 TCCTCAATCGTCACTCATGACAATGACATTTTCGTACGATTTTACCAATGGATCAAGAG 902
Db 81 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 100
Qy 903 TTTACTACCGTGTAAATAATCGGAAACAAGCTTATAGGATCAATAAAAAATCTGCTGTG 962
Db 101 PheThrTyrHisIleLysAspArgGluGlnAlaTyrGlyIleAsnLysLysSerGlyGln 120
Qy 963 AATGAAGAATAAACCACTGAC 986
Db 121 GluGluLysThrAsnAsnThrAsp 128

RESULT 13
S77680
streptokinase A (EC 3.4.-.-) (allele 19) - Streptococcus pyogenes (fragment)
C:Species: Streptococcus pyogenes
C>Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004
C:Accession: S77688
R:Kapur, V.; Kanjilal, S.; Hamrick, M.R.; Li, L.L.; Whittam, T.S.; Sawyer, S.A.; Musser,
Mol. Microbiol. 16, 509-519, 1995
A:Title: Molecular population genetic analysis of the streptokinase gene of Streptococcus
A:Reference number: S77671; MUID:96037795; PMID:7565111
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-128 <KAP>
A:Cross-references: UNIPROT:Q54695; UNIPARC:UPI000008B8DC; EMBL:U25870; NID:9818942; PII
A:Experimental source: strain ET76/M72
C:Genetics:
A:Gene: aka
C:Superfamily: streptokinase
C:Keywords: hydrolase; plasminogen activator; virulence

Alignment Scores:
Pred. No.: 5,29e-30 Length: 128
Score: 557.00 Matches: 107
Percent Similarity: 89.84% Conservative: 8
Best Local Similarity: 83.59% Mismatches: 13
Query Match: 18.95% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-10 (1-1661) x S77688 (1-128)

Qy 603 GTGCGGTTAGACCATATAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 662
Db 1 ValArgValArgProTyrLysGluLysProIleGlnThrProAlaLysSerValAspVal 20
Qy 663 GAATATACCTGACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCAAGTCTCAAA 722
Db 21 ArgTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProValLeuLys 40
Qy 723 GATACCTAAGCTATTGAAACCACTAGCTATCGGTGACACCATCATCATCTCAAGAATTACTA 782
Db 41 AsnThrLysLeuLysThrLeuAlaIleGlyThrValThrSerGlnGluLeu 60
Qy 783 GCTCAAGCACAAGCATTTTAAACAAAACCAACCCAGCTATACGATTATGAACGTGAC 842
Db 61 AlaGlnAlaGlnSerIleLeuAsnGluSerHisProAspTyrThrIleTyrGluArgAsp 80
Qy 843 TCCTCAATCGTCACTCATGACAATGACATTTTCGTACGATTTTACCAATGGATCAAGAG 902
Db 81 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 100
Qy 903 TTTACTACCGTGTAAATAATCGGAAACAAGCTTATAGGATCAATAAAAAATCTGCTGTG 962
Db 101 PheThrTyrHisIleLysAspArgGluGlnAlaTyrGlyIleAsnLysLysSerGlyGln 120
Qy 963 AATGAAGAATAAACCACTGAC 986
Db 121 GluGluLysThrAsnAsnThrAsp 128

RESULT 14
S77679
streptokinase A (EC 3.4.-.-) (allele 10) - Streptococcus pyogenes (fragment)
C:Species: Streptococcus pyogenes
C>Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004
C:Accession: S77679
R:Kapur, V.; Kanjilal, S.; Hamrick, M.R.; Li, L.L.; Whittam, T.S.; Sawyer, S.A.; Musser,
Mol. Microbiol. 16, 509-519, 1995
A:Title: Molecular population genetic analysis of the streptokinase gene of Streptococcus
A:Reference number: S77671; MUID:96037795; PMID:7565111
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-128 <KAP>
A:Cross-references: UNIPROT:Q54686; UNIPARC:UPI000008B8DF; EMBL:U25861; NID:9818924; PII
A:Experimental source: strain ET50/M43
C:Genetics:
A:Gene: aka
C:Superfamily: streptokinase
C:Keywords: hydrolase; plasminogen activator; virulence

Alignment Scores:
Pred. No.: 1.15e-29 Length: 128
Score: 552.00 Matches: 106
Percent Similarity: 89.06% Conservative: 8
Best Local Similarity: 82.81% Mismatches: 14
Query Match: 18.78% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-10 (1-1661) x S77679 (1-128)

Qy 603 GTGCGGTTAGACCATATAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 662
Db 1 ValArgValArgProTyrLysGluLysProIleGlnThrProAlaLysSerValAspVal 20
Qy 663 GAATATACCTGACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCAAGTCTCAAA 722
Db 21 ArgTyrAlaValGlnPheThrProLeuAsnProAspAspPheThrProValLeuLys 40
Qy 723 GATACCTAAGCTATTGAAACCACTAGCTATCGGTGACACCATCATCATCTCAAGAATTACTA 782
Db 41 AspThrLysLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 60
Qy 783 GCTCAAGCACAAGCATTTTAAACAAAACCAACCCAGCTATACGATTATGAACGTGAC 842
Db 61 AlaGlnAlaGlnSerIleLeuLysGluSerHisProAsnTyrThrIleHisGluArgAsp 80
Qy 843 TCCTCAATCGTCACTCATGACAATGACATTTTCGTACGATTTTACCAATGGATCAAGAG 902
Db 81 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProThrAspGlnGlu 100
Qy 903 TTTACTACCGTGTAAATAATCGGAAACAAGCTTATAGGATCAATAAAAAATCTGCTGTG 962
Db 101 PheThrTyrHisValLysAsnArgGluGlnAlaTyrLysAlaAsnSerLysThrAspIle 120
Qy 963 AATGAAGAATAAACCACTGAC 986

```

```
Db 121 LysGluLysThrAsnAsnThrAsp 128
||||: |||||||
RESULT 15
A43908
Fibronectin - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 12-Feb-1999
C:Accession: A43908
R:Desimone, D.W.; Norton, P.A.; Hynes, R.O.
Dev. Biol. 149, 357-369, 1992
A:Title: Identification and characterization of alternatively spliced fibronectin mRNAs
A:Reference number: A43908; MUID:921111942; PMID:1730390
A:Accession: A43908
A>Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-2481 <DES>
A:Cross-references: UNIPARC:UPI000177AE7; GB:M77820
A:Note: sequence extracted from NCBI backbone (NCBIP:77473)
C:Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II repe
C:Keywords: duplication; extracellular matrix; glycoprotein; heterodimer
F:55-90/Domain: fibronectin type I repeat homology <1F1>
F:100-138/Domain: fibronectin type I repeat homology <1F2>
F:144-182/Domain: fibronectin type I repeat homology <1F3>
F:189-228/Domain: fibronectin type I repeat homology <1F4>
F:234-273/Domain: fibronectin type I repeat homology <1F5>
F:309-343/Domain: fibronectin type I repeat homology <1F6>
F:361-402/Domain: fibronectin type II repeat homology <2F1>
F:421-462/Domain: fibronectin type II repeat homology <2F2>
F:471-509/Domain: fibronectin type I repeat homology <1F7>
F:519-556/Domain: fibronectin type I repeat homology <1F8>
F:562-600/Domain: fibronectin type I repeat homology <1F9>
F:610-693/Domain: fibronectin type III repeat homology <FN3A>
F:719-801/Domain: fibronectin type III repeat homology <FN3B>
F:810-891/Domain: fibronectin type III repeat homology <FN3C>
F:906-988/Domain: fibronectin type III repeat homology <FN3D>
F:996-1077/Domain: fibronectin type III repeat homology <FN3E>
F:1086-1165/Domain: fibronectin type III repeat homology <FN3F>
F:1173-1258/Domain: fibronectin type III repeat homology <FN3G>
F:1266-1349/Domain: fibronectin type III repeat homology <FN3H>
F:1357-1440/Domain: fibronectin type III repeat homology <FN3I>
F:1448-1530/Domain: fibronectin type III repeat homology <FN3J>
F:1538-1620/Domain: fibronectin type III repeat homology <FN3K>
F:1615-1617/Region: cell attachment (R-G-D) motif
F:1632-1714/Domain: fibronectin type III repeat homology <FN3L>
F:1722-1804/Domain: fibronectin type III repeat homology <FN3M>
F:1812-1894/Domain: fibronectin type III repeat homology <FN3N>
F:1904-1985/Domain: fibronectin type III repeat homology <FN3O>
F:1993-2075/Domain: fibronectin type III repeat homology <FN3P>
F:2197-2277/Domain: fibronectin type III repeat homology <FN3Q>
F:2301-2340/Domain: fibronectin type I repeat homology <1F10>
F:2346-2383/Domain: fibronectin type I repeat homology <1F11>
F:2390-2425/Domain: fibronectin type I repeat homology <1F12>
F:55-81,79-90,100-128,126-138,144-172,170-182,189-218,216-228,234-263,261-273,309-336,33
2373,2371-2383,2390-2416,2414-2425/Disulfide bonds: #status predicted
F:2459/disulfide bonds: interchain (to 2463) #status predicted
F:2463/disulfide bonds: interchain (to 2459) #status predicted

Alignment Scores:
Pred. No.: 7,87e-26 Length: 2481
Score: 496.00 Matches: 89
Percent Similarity: 78.12% Conservative: 11
Best Local Similarity: 69.53% Mismatches: 26
Query Match: 16.88% Indels: 2
DB: 2 Gaps: 1

US-09-940-235-10 (1-1661) x A43908 (1-2481)
```

```
Qy 1275 ATAACCGTTTATATGGCAGCGACCCGAGGAGGAGAAATGCTAGCTATCATTTAGCCGGT 1334
||||: |||
Db 15 LeuSerValValLeuArgAlaAlaProSerAlaThrSerLysLysArgGlnAlaGln 34
Qy 1335 GGTGCTAGCGCCAGCAAAATGTTTCAGCCGCCAGTCCCGGTGGCTGCTCAGTCAAGCAAG 1394
```

```
Db 35 GlnGlnGlnGlnGlnValValGlnProGlnGly-----ThrGlnAsnHisGln 52
||||: |||||||
Qy 1395 CCGGTTTGTATGACAAATGGAACACTATCAGATAAATCAACAGTGGGAGCGACCTAC 1454
||||: |||||||
Db 53 LysGlyCysTyrAspAsnGlyLysTyrTyrGlnIleAsnGlnGlnTrpGluArgThrTyr 72
||||: |||||||
Qy 1455 CTAGGTAATCGTTTGGTTTCTACTTGTATGAGAGAACCGAGGTTTAACTGCGAAAGT 1514
||||: |||||||
Db 73 LeuGlyAsnThrLeuValCysThrCysTyrGlyGlyArgGlyPheAsnCysGluSer 92
||||: |||||||
Qy 1515 AAACCTGAAGCTGAAGAGACTTGCTTTGACAAAGTACATCGGGAACACTTACCGAGTGGGT 1574
||||: |||||||
Db 93 LysProGluSerGluGluThrCysPheAspLysTyrThrGlyValSerTyrArgValGly 112
||||: |||||||
Qy 1575 GACACTTATGAGCGTCTCTAAAGACTCCATGATCTGGGACTGTACCTGCATCGGGGCTGGG 1634
||||: |||||||
Db 113 GluThrTyrGluArgProLysAspAsnMetIleTrpAspCysThrCysIleGlyAlaGly 132
||||: |||||||
Qy 1635 CGAGGGAGAGATAAGCTGTACCATC 1658
||||: |||||||
Db 133 ArgGlyArgIleSerCysThrIle 140
||||: |||||||
```

Search completed: January 28, 2006, 02:37:37

Job time : 40.4546 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 28, 2006, 01:48:09 ; Search time 179.003 Seconds
(without alignments)
13093.422 Million cell updates/sec

Title: US-09-940-235-10
Perfect score: 2939
Sequence: 1 gcaacccgccagctagcc.....gaataagctgtaccatctaa 1661

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 4332886

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_epool_p/US09940235/runat_27012006.144218.27563/app.query.fasta_1.7708
-DB=UniProt -QFMT=fastan -SUFFIX=rup -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptio -NORM=ext -HRAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09940235 @CGN_1_1_1359 @runat_27012006.144218.27563 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : UniProt_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1984	67.5	440	1 STRP_STRQ	P00779 streptococc
2	1955	66.5	414	2 Q53284_STRQ	Q53284 streptococc
3	1943	66.1	436	2 Q6UK57_STRQ	Q6UK57 streptococc
4	1942	66.1	440	1 STRP_STRS1	P10519 streptococc
5	1814	61.7	440	2 Q7X0Y1_STRPY	Q7X0Y1 streptococc
6	1802	61.3	440	2 Q7X0Y8_STRPY	Q7X0Y8 streptococc
7	1802	61.3	440	2 Q8X5R8_STRP3	Q8X5R8 streptococc
8	1794	61.0	440	1 STRP_STRPY	P10520 streptococc
9	1794	61.0	440	2 Q532X6_STRPY	Q532X6 streptococc
10	1787	60.8	440	2 Q7X0Y2_STRPY	Q7X0Y2 streptococc
11	1758	59.8	440	2 Q7X0Y7_STRPY	Q7X0Y7 streptococc
12	1758	59.8	440	2 Q5X9T6_STRP6	Q5X9T6 streptococc
13	1755	59.7	440	2 Q8NZA6_STRP8	Q8NZA6 streptococc
14	1723	58.6	440	2 Q7X0Y3_STRPY	Q7X0Y3 streptococc
15	1715	58.4	440	2 Q7X0Y0_STRPY	Q7X0Y0 streptococc
16	1715	58.4	440	2 Q7X0Y5_STRPY	Q7X0Y5 streptococc

17	1708	58.1	440	2	Q7X0X8_STRPY	Q7X0X8 streptococc
18	1686	57.4	440	2	Q7X0X9_STRPY	Q7X0X9 streptococc
19	1677	57.1	440	2	Q7X0X7_STRPY	Q7X0X7 streptococc
20	1670	56.8	440	2	Q57391_STRPY	Q57391 streptococc
21	1657	56.4	440	2	Q7X0Y6_STRPY	Q7X0Y6 streptococc
22	1652	56.2	440	1	STRQ_STRPY	P96471 streptococc
23	1624	55.3	432	2	Q7X0Y4_STRPY	Q7X0Y4 streptococc
24	728	24.8	141	2	Q7X0X2_STRQ	Q7X0X2 streptococc
25	727	24.7	141	2	Q7X0X3_STRQ	Q7X0X3 streptococc
26	724	24.6	141	2	Q7X0W1_STRQ	Q7X0W1 streptococc
27	723	24.6	141	2	Q7WS87_STRQ	Q7WS87 streptococc
28	717	24.4	141	2	Q7X0W3_STRQ	Q7X0W3 streptococc
29	717	24.4	141	2	Q7X0X4_STRQ	Q7X0X4 streptococc
30	715	24.3	141	2	Q7X0W5_STRQ	Q7X0W5 streptococc
31	714	24.3	141	2	Q7X0X5_STRQ	Q7X0X5 streptococc
32	712	24.2	141	2	Q7X0X6_STRQ	Q7X0X6 streptococc
33	711	24.2	141	2	Q7X0W2_STRQ	Q7X0W2 streptococc
34	711	24.2	141	2	Q7X0W9_STRQ	Q7X0W9 streptococc
35	710	24.2	141	2	Q7X0W6_STRQ	Q7X0W6 streptococc
36	707	24.1	141	2	Q7X0W0_STRQ	Q7X0W0 streptococc
37	705	24.0	141	2	Q7X0V5_STRPY	Q7X0V5 streptococc
38	704	24.0	141	2	Q7X0X0_STRQ	Q7X0X0 streptococc
39	701	23.9	141	2	Q7X0W7_STRQ	Q7X0W7 streptococc
40	697	23.7	141	2	Q7X0W4_STRQ	Q7X0W4 streptococc
41	696	23.7	141	2	Q7X0W8_STRQ	Q7X0W8 streptococc
42	693	23.6	141	2	Q7X0X1_STRQ	Q7X0X1 streptococc
43	690	23.5	141	2	Q7X0R2_STRPY	Q7X0R2 streptococc
44	686	23.3	141	2	Q7X0S4_STRPY	Q7X0S4 streptococc
45	684	23.3	141	2	Q7X0S0_STRPY	Q7X0S0 streptococc

ALIGNMENTS

RESULT 1

ID	STRP_STRQ	STANDARD;	PRT;	440 AA.
AC	P00779;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	01-APR-1988 (Rel. 07, Last sequence update)			
DT	10-MAY-2005 (Rel. 47, Last annotation update)			
DE	Streptokinase C precursor.			
GN	Name=skc;			
OS	Streptococcus equisimilis.			
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;			
OC	Streptococcus.			
OX	NCBI_TaxID=119602;			
RN	[1]			
RP	NUCLEOTIDE SEQUENCE.			
RC	STRAIN=H46A;			
RA	MEDLINE=85232082; PubMed=2989113; DOI=10.1016/0378-1119(85)90145-3;			
RA	Malke H., Roe B.A., Ferretti J.J.;			
RT	"Nucleotide sequence of the streptokinase gene from Streptococcus equisimilis H46A.";			
RL	Gene 34:357-362(1985).			
RN	[2]			
RP	PROTEIN SEQUENCE OF 27-440.			
RA	MEDLINE=83127125; PubMed=6760891;			
RA	Jackson K.W., Tang J.;			
RT	"Complete amino acid sequence of streptokinase and its homology with serine proteases.";			
RL	Biochemistry 21:6620-6625(1982).			
CC	-!- FUNCTION: This protein is not a protease, but it activates plasminogen by complexing with it. As a potential virulence factor, it is thought to prevent the formation of effective fibrin barriers around the site of infection, thereby contributing to the invasiveness of the cells.			
CC	-----			
CC	This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.			
CC	-----			

DR EMBL; K02986; AAA26974.1; -: Genomic DNA.
 DR EMBL; X72832; CAA51351.1; -: Genomic DNA.
 DR PIR; A09667; BZSO.
 DR PIR; A22801; A22801.
 DR PDB; 1BML; X-ray; C/D=38-399.
 DR PDB; 1L4D; X-ray; B=40-173.
 DR PDB; 1L4Z; X-ray; B=27-173.
 DR PDB; 1OOR; X-ray; A/B/C/D=177-314.
 DR SMR; P00779; 38-398.
 DR InterPro; IPR004093; Staphylokinase.
 DR InterPro; IPR008124; Streptokinase.
 DR Pfam; PF02821; Staphylokinase; 3.
 DR PRINTS; PR01753; STREPKINASE.
 KW 3D-structure; Direct protein sequencing; Plasminogen activation;
 KW Signal; Virulence. 26
 FT SIGNAL 1
 FT CHAIN 27 440 Streptokinase C.
 FT VARIANT 195 195 L -> D.
 FT VARIANT 207 207 D -> L.
 FT CONFLICT 298 300 EKV -> LEVK (in Ref. 2).
 FT CONFLICT 438 438 N -> D (in Ref. 2).
 FT STRAND 180 180
 FT STRAND 184 194
 FT STRAND 199 200
 FT TURN 203 204
 FT STRAND 205 205
 FT STRAND 209 214
 FT TURN 216 217
 FT STRAND 219 221
 FT HELIX 222 236
 FT TURN 238 239
 FT STRAND 240 252
 FT TURN 253 254
 FT STRAND 259 260
 FT STRAND 268 270
 FT STRAND 275 275
 FT STRAND 278 280
 FT TURN 282 284
 FT STRAND 287 289
 FT TURN 292 304
 FT TURN 305 306
 SQ SEQUENCE 440 AA; 50140 MW; 8FC1F22648ACC77A CRC64;

Alignment Scores:

Pred. No.: 2,586-128 Length: 440
 Score: 1984.00 Matches: 383
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 67.51% Indels: 0
 DB: 1 Gaps: 0

US-09-940-235-10 (1-1661) x STRP_STREQ (1-440)

QY 183 ATTGCTGGACCTGAGTGGCTGTAGACCTGCTCAACAAACAGCCAAATGGTGT 242
 DB 27 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 46
 QY 243 AGCGTGTGCTGCTGTTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTGAATC 302
 DB 47 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 66
 QY 303 GATCTAACTACGACCTGCTCATGGAGGAAGACAGACAGGCTTAAGTCCAAATCA 362
 DB 67 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 86
 QY 363 AAACCATTTGCTGATGATGCGCGGATGCACATAAACTTGAGAAAGCTGACTTACTA 422
 DB 87 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlnLysAlaAspLeuLeu 106
 QY 423 AAGGCTATTCAAGAACTTGTACGCTACGCTCCACAGTAACGACGACTACTTTGAGGTC 482
 DB 107 LysAlaIleGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 126

QY 483 ATTGATTTTCAAGCGATGCAACATTTACTGATCGAAACGCAAGGTCTACTTTGCTGAC 542
 DB 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146
 QY 543 AAAGATGGTTGCGTAACTTTGCCGACCCAACTGTCTCAAGAAATTTTGTAAAGCGGACAT 602
 DB 147 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 166
 QY 603 GTGCGGTTAGACCATATAAAGAAAACCAATCAACCAAGGGAATCTGTTGATGTG 662
 DB 167 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 186
 QY 663 GAATATACTGTACAGTTTACTCCCTTAAACCCGTGATGACGATTTTACAGCAGGTCTCAA 722
 DB 187 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 206
 QY 723 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTA 782
 DB 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 226
 QY 783 GCTCAAGCACAAAGCATTTTAAACAAAACCCAGGCTATAGCATTTATGAACGTGAC 842
 DB 227 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 246
 QY 843 TCCTCAATCGTCACTCATGACATTTCCGTACGATTTTACCAATGGATCAAGAG 902
 DB 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
 QY 903 TTTTACTTACCGTGTATAAAATCGGGAAACAGCTTTATAGGATCAATAAAAAATCTGGTCTG 962
 DB 267 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 286
 QY 963 AATGAAGAAATAAACCACTGACCTGATCTCTGAGAAATATTACGTCTTAAAAAGGG 1022
 DB 287 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 306
 QY 1023 GAAAAGCGTATGATCCCTTTGATCGCAGTCACCTGAAACTGTTTACCACCAATAGTGT 1082
 DB 307 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326
 QY 1083 GATGTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAAGTAA 1142
 DB 327 AspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
 QY 1143 TTAGACTTCAGAGATTTATACGATCTCTCGTGAAGGCTAACTACTCTTCAACAATCTC 1202
 DB 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
 QY 1203 GATGCTTTTGGTATTATGGACTATACCTTAACCTGGAAGAGTAGGAGGATATCAGATGAC 1262
 DB 367 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 386
 QY 1263 ACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAAGGAGAGAATGCTAGTCTAT 1322
 DB 387 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 406
 QY 1323 CATTTAGCC 1331
 DB 407 HisLeuAla 409
 RESULT 2
 Q53284 STREQ
 ID Q53284 STREQ PRELIMINARY; PRT; 414 AA.
 AC Q53284;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE SKC-2.
 OS Streptococcus equisimilis.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=119602;
 RN [1]

QY 183 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCCAATTGGTTGTT 242

Db 23 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal 42
 QY AGCGTTGCTGTTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTGAATC 302
 Db 43 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPheGluIle 62
 QY 303 GATCTAACATCACGACCTGCTCATGAGGAGAAACAGACAGCAAGCTTAAGTCCAAATCA 362
 Db 63 AspLeuThrSerGlnProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 82
 QY 363 AAACATTGCTGATGATGCGCGATGTCATATAAATTGAGAAAGCTGATCTACTA 422
 Db 83 LysProPheAlaThrAspSerGlyAlaMetProHisLeuLeuGluLysAlaAspLeuLeu 102
 QY 423 AAGGCTATTCAAGACAAATTGATCGCTTAAGTCCACAGTAACGACGACTACTTTGAGTTC 482
 Db 103 LysAlaIleGlnGluLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 122
 QY 483 ATTGATTTTGAAGCGATGCAACCATTTACTGATCGAAACGCAAGCTCTACTTTGCTGAC 542
 Db 123 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 142
 QY 543 AAAGATGTTTGGTAACTTCCGACCCCAACCTGTCCTCAAGAAATTTTGTCTAAGCGGACAT 602
 Db 143 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuLysGlyHis 162
 QY 603 GTGCGCTAGACCATATAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTC 562
 Db 163 ValArgValArgProTyrLysGluLysProValGlnAsnGlnAlaLysSerValAspVal 182
 QY 663 GAATATATCTGATGATTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTCAA 722
 Db 183 GluThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 202
 QY 723 GATCTAAGCTATTGAAACACTAGCTATCGGTGACCATCACATCTCAAGAATTTACTA 782
 Db 203 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 222
 QY 783 GCTCAAGCACAAAGCATTTTAAACAAACCCAGGCTATACGATTTATGACGTCGAC 842
 Db 223 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp 242
 QY 843 TCCTCAATCGTCACATCATGACATGATTTTCGTCAGCATTTTACCATGATCATCAAG 902
 Db 243 SerSerIleValThrHisAspLysAspIlePheArgThrIleLeuProMetAspGlnGlu 262
 QY 903 TTTACTTACCGTGTAAAAATCGGAAACAAAGCTTATAGGATCAATAAAAAATCTGCTCG 962
 Db 263 PheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsnLysLysSerGlyLeu 282
 QY 963 AATGAGAAATAACAAACACTGACCTGATCTCTGAGAAATATATCGTCTTAAAAAGGG 1022
 Db 283 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 302
 QY 1023 GAAAGCCGATGATCCCTTTGATCGCAGTCATCTGAACTGTTTCCATCAATACGTT 1082
 Db 303 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 322
 QY 1083 GATGTCGATACCAACGAATTGCTAAAAAGTGAAGAGCTCTTAAACAGTAGCGCAACTAAC 1142
 Db 323 AspValAsnThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 342
 QY 1143 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACAACTCTC 1202
 Db 343 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 362
 QY 1203 GATGCTTTTGTATTTATGACTATACCTTAACCTGAAAAAGTAGAGATATACGATGAC 1262
 Db 363 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 382
 QY 1263 ACCAACCGTATCATAAACGTTTATATGGCAAGCGACCCGAAAGGAGAGATGCTAGCTAT 1322

Db 383 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 402
 QY 1323 CATTTAGCC 1331
 Db 403 HisLeuAla 405
 RESULT 4
 STRP_STRS1
 ID STRP_STRS1 STANDARD; PRT; 440 AA.
 AC P10519;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Streptokinase G precursor.
 GN Name=skg;
 OS Streptococcus sp. (strain 19909).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 ON NCBI_TaxID=69017;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
 RA MEDLINE=89160285; PubMed=2922269;
 RX Walter F., Siegel M., Malke H.;
 RT "Nucleotide sequence of the streptokinase gene from a group-G
 Streptococcus";
 RL Nucleic Acids Res. 17:1262-1262(1989).
 CC -!- FUNCTION: This protein is not a protease, but it activates
 plasminogen by complexing with it. As a potential virulence
 factor, it is thought to prevent the formation of effective fibrin
 barriers around the site of infection, thereby contributing to the
 invasiveness of the cells.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use as long as its content is in no way modified and this statement is not
 removed.
 CC -----
 CC EMBL: X13400; CAA31766.1; -; Genomic_DNA.
 DR PIR: S02723; S02723.
 DR HSP: P00779; 1QOR.
 DR SMR: P10519; 63-398.
 DR InterPro: IPR004093; Staphylokinase.
 DR InterPro: IPR008124; Streptokinase.
 DR Pfam: PF02821; Staphylokinase; 3.
 DR PRINTS: PR01753; STREPKINASE.
 KW Plasminogen activation; Signal; Virulence.
 FT SIGNAL 1 26
 FT CHAIN 27 440 Streptokinase G.
 SQ SEQUENCE 440 AA; 50199 MW; 5521P8825F81B6EA CRC64;
 Alignment Scores:
 Pred. No.: 2,03e-125 Length: 440
 Score: 1942.00 Matches: 375
 Percent Similarity: 98.43% Conservativeness: 2
 Best Local Similarity: 97.91% Mismatches: 6
 Query Match: 66.08% Indels: 0
 DB: Gaps: 0
 US-09-940-235-10 (1-1661) x STRP_STRS1 (1-440)
 QY 183 ATTCGTCGACCTGAGTGGCTGTAGACCGTCTGTCATCTGTCACACAGCCAAATGGTTGTT 242
 Db 27 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal 46
 QY 243 AGCGTTGCTGTTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTGAATC 302
 Db 47 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPheGluIle 66
 QY 303 GATCTAACATCACGACCTGCTCATGAGGAGAAACAGACAGCAAGCTTAAAGTCCAAATCA 362
 Db 67 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 86


```
QY 363 AAACCAATTTGCTACTGATAGTGGCGCGATGTCAATAACTTGAGAAAGCTGACTTACTA 422
|||
Db 87 LysLeuPheAlaThrAspSerGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu 106
|||
QY 423 AAGGCTATTCAAGAACAAATGATCGCTAAAGTCCACAGTACGACGACTACTTTGAGGTC 482
|||
Db 107 LysAlaIleGlnGluLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 126
|||
QY 483 ATTGATTTTTCGAACGATGCAACATTTACTGATCGAAACGCAAGGCTCTACTTTGCTGAC 542
|||
Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146
|||
QY 543 AAGATGTTTCGGTAACCTTGGCGACCCCAACCTGTCCAAGATTTTTCCTAAGCGACAT 602
|||
Db 147 LysAspGlySerValThrLeuProIleGlnProValGlnGluPheLeuLeuLysGlyHis 166
|||
QY 603 GTGCGGTTAGACCATATAAGAAACCAACCAATACAAACCAACGGAATCTGTTGATGTG 662
|||
Db 167 ValArgValArgProTyrLysGluLysProValGlnAsnGlnAlaLysSerValAspVal 186
|||
QY 663 GAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCAAGGTCCTCAA 722
|||
Db 187 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProAlaLeuLys 206
|||
QY 723 GATCTAAGCTATTGAAACACTAGTATCGGTGACACCACTCATCATCTCAAGAAATTACTA 782
|||
Db 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 226
|||
QY 783 GCTCAAGCACAAGCAATTTAAACAAACCAACCCAGGCTATACGATTTATGAACGTGAC 842
|||
Db 227 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 246
|||
QY 843 TCTCTCAATGCTCACTCATGACATGACATTTTCGTACGATTTTACCATGATCAAGAG 902
|||
Db 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
|||
QY 903 TTTACTTACCGTGTAAAAATCGGGAACAGCTTTATAGGATCAATAAAAAATCTGGCTG 962
|||
Db 267 PheThrTyrHisValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 286
|||
QY 963 AATGAAGAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1022
|||
Db 287 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 306
|||
QY 1023 GAAAGCCGTATGATCCCTTTGATCGCAGTCACCTTGAACCTGTTCAACATCAATAGTT 1082
|||
Db 307 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326
|||
QY 1083 GATGTCGATACCAACCAATTTGCTAAAGTGACAGCTCTTAAACAGCTAGCGACGTAAC 1142
|||
Db 327 AspValAsnThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
|||
QY 1143 TTAGACTTCAGAGATTATACGATCCCTCGTGATAAGCTAAACTCTCTACAACAATCTC 1202
|||
Db 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
|||
QY 1203 GATGCTTTTGGTATTATGACTATACCTTAACCTGGAAGATAGAGGATATACGATGAC 1262
|||
Db 367 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 386
|||
QY 1263 ACCAACCGTATCATACCGTTTATATGGCAACGCGACCGAAGGAGATGCTAGCTAT 1322
|||
Db 387 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 406
|||
QY 1323 CATTTAGCC 1331
|||
Db 407 HisLeuAla 409
|||
```

RESULT 5

Q7X0Y1_STRPY

ID Q7X0Y1_STRPY PRELIMINARY; PRT; 440 AA.

AC Q7X0Y1.

DT 01-OCT-2003 (TrEMBLrel. 25, Created)

```
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Streptokinase.
GN Name=ska;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]_
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=D488;
RX PubMed=14579231; DOI=10.1128/JB.186.1.110-121.2004;
RA Kalia A., Bessen D.E.;
RT "Natural selection and evolution of streptococcal virulence genes
involved in tissue-specific adaptations.";
RL J. Bacteriol. 186:110-121(2004).
DR EMBL; AY234137; AAP39957.1; -; Genomic_DNA.
DR HSBP; Q53284; 1C4P.
DR SMR; Q7X0Y1; 63-398.
DR GO; GO:0016301; P:kinase activity; IEA.
DR GO; GO:0008243; P:plasminogen activator activity; IEA.
DR GO; GO:0005515; P:protein binding; IEA.
DR InterPro; IPR004093; Staphylokinase.
DR InterPro; IPR008124; Streptokinase.
DR Pfam; PF02821; Staphylokinase; 3.
DR PRINTS; PR01753; STREPKINASE.
DR Kinase.
SQ SEQUENCE 440 AA; 49903 MW; CC4B659647043BAC CRC64;

Alignment Scores:
Pred. No.: 1,36e-116 Length: 440
Score: 1814.00 Matches: 347
Percent Similarity: 95.04% Conservative: 17
Best Local Similarity: 90.60% Mismatches: 19
Query Match: 61.72% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-10 (1-1661) x Q7X0Y1_STRPY (1-440)
QY 183 ATTCTCGACCTGAGTGGCTGTAGACGCTCCATCTGTCAACAAACAGCAATGTTGTTT 242
|||
Db 27 IleAlaGlyTyrGlyTyrLeuProAspArgProIleAsnAsnSerGlnLeuValVal 46
|||
QY 243 AGCCTTCTGCTACTGTTGAGGCGACCAATCAACACATAGCTCTTAATTTTTTGAATC 302
|||
Db 47 SerMetAlaGlyIleValGluGlyThrAspLysLysValPheIleAsnPheGluIle 66
|||
QY 303 GATCTAATCACTCAGCACTGCTCATGGAGGAAGACAGACGAGCTTAAGTCCAAATCA 362
|||
Db 67 AspLeuThrSerGlnProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 86
|||
QY 363 AAACATTTTGTCTACTGATAGTGGCGCGATGTCAATAACTTGAGAAAGCTGACTTACTA 422
|||
Db 87 LysProPheAlaThrAspAsnGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu 106
|||
QY 423 AAGCTATTCAAGAACAAATGATCGCTAAAGTCCACAGTACGACGACTACTTTGAGGTC 482
|||
Db 107 LysAlaIleGlnLysGlnLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal 126
|||
QY 483 ATTGATTTTTCGAACGATGCAACATTTACTGATCGAAACGCAAGGCTCTACTTTGCTGAC 542
|||
Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyGlnValTyrPheAlaAsp 146
|||
QY 543 AAGATGTTTCGGTAACCTTGGCGACCCCAACCTGTCCAAGATTTTTCCTAAGCGACAT 602
|||
Db 147 LysAspGlySerValThrLeuProIleGlnProValGlnGluPheLeuLeuLysGlyHis 166
|||
QY 603 GTGCGGTTAGACCATATAAGAAACCAACCAATACAAACCAACGGAATCTGTTGATGTG 662
|||
Db 167 ValArgValArgProTyrLysGluLysProValGlnAsnGlnAlaLysSerValAspVal 186
|||
QY 663 GAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCAAGGTCCTCAA 722
|||
```

Db 187 LysThrValGlnPheThrProLeuAenProAaspAaspPheArgProGlyLeuLys 206
QY 723 GATACCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAACTACTA 782
Db 207 AspThrLysLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 226
QY 783 GCTCAAGCAAAAGCAATTTAAACAAACCCAGGCTATACGATTTATGATCAACGTGAC 842
Db 227 AlaGlnAlaGlnSerIleLeuAenLysThrHisProGlyThrIleThrLysGluArgAsp 246
QY 843 TCCTCAATCGTCACTCATGACATGACATTTCCGTACGATTTTACCAATGATGATCAAGAG 902
Db 247 SerSerIleValThrHisAspAenAspIlePheArgThrIleLeuProMetAspGlnGlu 266
QY 903 TTTACTTACCGTGTAAATAATCGGAACCAAGCTTATAGATCAATAAAATAATCTGTCTG 962
Db 267 PheThrTyArgValLysAspArgGluGlnAlaTyGlyLeuLysLysSerGlyLeu 286
QY 963 AATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGCTCTTAAAAAGGG 1022
Db 287 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyThrIleLeuLysLysGly 306
QY 1023 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAATGAACTGTTCAACATCAATACGTT 1082
Db 307 GluSerProTyAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyVal 326
QY 1083 GATGTCGATACCAAGAAATGCTTAAAGTGAAGCAGCTCTTAACAGCTAGCGAACCTAAC 1142
Db 327 AspValAsnThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
QY 1143 TTAGACTTCAGAGATTTATACGATCTCTGATGATAAGGCTAAACTACTCTCAACAATCTC 1202
Db 347 LeuAspPheArgAspLeuTyAspProArgAspLysAlaLysLeuLeuTyAsnAsnLeu 366
QY 1203 GATGCTTTGGTATTATGACATATACCTTAACCTGAAAGTGAAGGATATCAACGATGAC 1262
Db 367 AspAlaPheAspIleMetAspTyThrLeuThrGlyLysValGluAspAsnHisAspLys 386
QY 1263 ACCACCGTATCATACCGTTTATATGCGCAGCGACCGACCGAGGAGATGCTACTAT 1322
Db 387 AsnAsnArgIleValThrValTyMetGlyLysArgProLysGlyAlaLysGlySerTy 406
QY 1323 CATTTAGCC 1331
Db 407 HisLeuAla 409

RESULT 6

Q7X0Y8 STRPY
ID Q7X0Y8 STRPY PRELIMINARY; PRT; 440 AA.
AC Q7X0Y8
DT 01-OCT-2003 (TREMBLrel. 25, Created)
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Streptokinase.
GN Name=ska;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=88-019;
RX PubMed=14679231; DOI=10.1128/JB.186.1.110-121.2004;
RA Kalia A., Bessen D.E.;
RT "Natural selection and evolution of streptococcal virulence genes
involved in tissue-specific adaptations.";
RL J. Bacteriol. 186:110-121(2004).
DR EMBL; AY234129; AAP39949.1; -; Genomic_DNA.
DR HSSP; Q53284; IC4P.
DR SMR; Q7X0Y8; 63-398.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008243; F:plasminogen activator activity; IEA.
DR GO; GO:0005515; F:protein binding; IEA.

DR InterPro; IPR004093; Staphylokinase.
DR InterPro; IPR008124; Streptokinase.
DR Pfam; PF02821; Staphylokinase; 3.
DR PRINTS; PR01753; STREPKINASE.
KW Kinase.
SQ SEQUENCE 440 AA; 49885 MW; 69DB44F4026E3975 CRC64;

Alignment Scores:

Pred. No.: 9.12e-116 Length: 440
Score: 1802.00 Matches: 347
Percent Similarity: 94.52% Conservative: 15
Best Local Similarity: 90.60% Mismatches: 21
Query Match: 61.31% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-10 (1-1661) x Q7X0Y8_STRPY (1-440)

QY 183 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTCTCAACACACGCCAATTTGTTT 242
Db 27 IleAlaGlyTyThrLeuProAspArgProValAsnAsnSerGlnLeuVal 46
QY 243 AGCGTTCCTGCTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAATTTTGAATC 302
Db 47 SerMetAlaGlyIleValGluGlyThrAspLysLysValPheIleAsnPheGluIle 66
QY 303 GATCTAACATCACGACCTGCTCATGGAGGAAAGACAGACGAAGCTTTAAGTCCAAATCA 362
Db 67 AspLeuThrSerGlnHisAlaHisGlyGlyThrGlnGlyLeuSerProLysSer 86
QY 363 AAACCAATTTGCTACTGATAGTGGCGGATGTCATAAACTTGAGAAAGCTGACTTACTA 422
Db 87 LysProPheAlaThrAspAsnGlyAlaMetProHisLysLeuGluLysAlaAspLeu 106
QY 423 AAGGCTATTCAAGAACAAATTTGATCGCTTAACGTCACAGTAAACGACGACTTTGAGTC 482
Db 107 LysAlaIleGlnLysGlnLeuIleAlaAsnValHisSerAsnAspGlyTyThrPheGluVal 126
QY 483 ATTGATTTTGCAGCGATGCAACCATTTACTGATGCGCAACCGCAAGCTCTACTTTGCTGAC 542
Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyThrPheAlaAsp 146
QY 543 AAAGATGTTTCCGTTAACTCCCTTAAACCCCTGATGACGATTTTCAGACCGGTCTCAA 602
Db 147 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 166
QY 603 GTCGCGTTAGACCATATAAAGAAAAACCAATACAAACCAACGCAATCTGTTGATGTG 662
Db 167 ValArgValArgProTyLysGluLysProValGlnAsnGlnAlaLysSerValAspVal 186
QY 663 GAATATACGTACAGTATTCTCCCTTAAACCCCTGATGACGATTTTCAGACCGGTCTCAA 722
Db 187 LysTyThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 206
QY 723 GATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTA 782
Db 207 AspThrLysLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 226
QY 783 GCTCAAGCAAAAGCAATTTTAAACAAACCCAGGCTTATACGATTTTATGATCAACGTGAC 842
Db 227 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyThrIleTyGluArgAsp 246
QY 843 TCCTCAATCGTCACTCATGACATGATTTTCGTACGATTTTACCAATGATGATCAAGAG 902
Db 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
QY 903 TTTACTTACCGTGTAAATAATCGGAACCAAGCTTATAGGATCAATAAAATAATCTGTCTG 962
Db 267 PheThrTyArgValLysAspArgGluGlnAlaTyGlyIleAsnLysLysSerGlyLeu 286
QY 963 AATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGCTCTTAAAAAGGG 1022
Db 287 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyThrIleLeuLysLysGly 306

```
QY 1023 GAAAGCCGATGATCCCTTTGATCGAGTCACCTTGAAACCTGTTACCACCAATGCTT 1082
Db |||||
QY 307 GluSerProTyAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326
Db |||||
QY 1083 GATGTCGATACCAACGAATTCCTAAAGATGACGACGCTCTTAACAGCTACGGAACGTAAC 1142
Db |||||
QY 327 AspValAsnThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
Db |||||
QY 1143 TTAGACTTCAGAGATTATACCATCCTCGTGATAAGCTAAACTACTCTCAACAATCTC 1202
Db |||||
QY 347 LeuAspPheArgAspLeuTyrAspProCysAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
Db |||||
QY 1203 GATGCTTTGGTATTATGACATATACCTTAACCTGGAAGATAGAGGATATACGATGAC 1262
Db |||||
QY 367 AspAlaPheAspIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspLys 386
Db |||||
QY 1263 ACCACCGTATCATACACCTTTATATGCGCAAGCGACCCGAGAGAGATGCTAGCTAT 1322
Db |||||
QY 387 AsnAsnArgIleValThrValTyrMetGlyLysArgProLysGlyAlaLysGlySerTyr 406
Db |||||
QY 1323 CATTTAGCC 1331
Db |||||
QY 407 HisLeuAla 409
Db |||||
RESULT 7
ID Q8K5R8_STRP3 PRELIMINARY; PRT; 440 AA.
AC Q8K5R8_079W73;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 13-SEP-2005 (TReMBLrel. 31, Last annotation update)
DE Streptokinase A.
GN Name=ska; OrderedLocusNames=SPs1700, SpyM3_1698;
OS Streptococcus pyogenes (serotype M3).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=301448;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MGA5315 / Serotype M3;
RX MEDLINE=22133808; PubMed=12122206; DOI=10.1073/pnas.152298499;
RA Beres S.B., Sylva G.L., Barbhan K.D., Lei B., Hoff J.S.,
RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,
RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
RA Schlievert P.M., Musser J.M.;
RT "Genome sequence of a serotype M3 strain of group A Streptococcus:
RT phage-encoded toxins, the high-virulence phenotype, and clone
RT emergence.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SSI-1 / Serotype M3;
RX MEDLINE=22683278; PubMed=12799345; DOI=10.1101/gr.1096703;
RA Nakagawa I., Kurokawa K., Yamashita A., Nakata M., Tomiyasu Y.,
RA Okahashi N., Kawabata S., Yamazaki K., Shiba T., Yasunaga T.,
RA Hayashi H., Hattori M., Hamada S.;
RT "Genome sequence of an M3 strain of Streptococcus pyogenes reveals a
RT large-scale genomic rearrangement in invasive strains and new insights
RT into phage evolution.";
RL Genome Res. 13:1042-1055(2003).
DR EMBL; AE014169; AAM80305.1; -; Genomic DNA.
DR EMBL; BA000034; BAC64795.1; -; Genomic DNA.
DR HSSP; P00779; IL4D.
DR SMR; Q8K5R8; 63-398.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008243; F:plasminogen activator activity; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR InterPro; IPR004093; Staphylokinase.
DR InterPro; IPR008124; Streptokinase.
DR Pfam; PF02821; Staphylokinase; 3.
DR PRINTS; PR01753; STREPKINASE.
KW Complete proteome; Kinase.
SEQUENCE 440 AA; 49911 MW; 7CCE44F4026E3975 CRC64;
```

Alignment Scores:

```
Pred. No.: 9.12e-116 Length: 440
Score: 1802.00 Matches: 347
Percent Similarity: 94.52% Conservatives: 15
Best Local Similarity: 90.60% Mismatches: 21
Query Match: 61.31% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-10 (1-1661) x Q8K5R8_STRP3 (1-440)

QY 183 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCCAATTGGTTGTT 242
Db |||||
QY 27 IleAlaGlyTyrGlyTyrLeuProAspArgProValAsnAsnSerGlnLeuValVal 46
Db |||||
QY 243 AGCGTCTCGTACTGTTGAGGGGACGAATCAAGACATAGCTTAAATTTTTTGAATC 302
Db |||||
QY 47 SerMetAlaGlyIleValGluGlyThrAspLysLysValPheIleAsnPheGluIle 66
Db |||||
QY 303 GATCTAACATCACGACCTGCTCATGGAGGAAAGACAGACAGCAAGCTTAAGTCCAAATCA 362
Db |||||
QY 67 AspLeuThrSerGlnHisAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 86
Db |||||
QY 363 AAACCATTTGCTACTGATAGTGGCGCATGTGCATATAAATTTGAGAAGCTGACTTACTA 422
Db |||||
QY 87 LysProPheAlaThrAspAsnGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu 106
Db |||||
QY 423 AAGCTATTCAAGAACAAATTGATCGCTAAAGTCCACAGTAAACGACGACTACTTTGAGTC 482
Db |||||
QY 107 LysAlaIleGlnLysGlnLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal 126
Db |||||
QY 483 ATTGATTTTCAAGCGATGCAACCATTTACTGATGATGAAACGCAAGGCTACTTTGCTGAC 542
Db |||||
QY 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146
Db |||||
QY 543 AAAGATGTTTCGGTAACTCTTCCGACCCCAACTCTCAAGAAATTTTTCCTAAGCGGACAT 602
Db |||||
QY 147 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 166
Db |||||
QY 603 GTGCGCTTAGACCATATAAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 662
Db |||||
QY 167 ValArgValArgProTyrLysGluLysProValGlnAsnGlnAlaLysSerValAspVal 186
Db |||||
QY 663 GAATATACTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTCAGACCGAGTCTCAA 722
Db |||||
QY 187 LysTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 206
Db |||||
QY 723 GATACTAAGCTATTGAAACACACTAGCTATCGGTGACACCATCACATCTCAAGAAATTA 782
Db |||||
QY 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 226
Db |||||
QY 783 GCTCAAGCACAAAGCAATTTTAAACAAACCAACCCAGCGCTATACGATTTATGAACGTGAC 842
Db |||||
QY 227 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp 246
Db |||||
QY 843 TCCTCAATCTCTACTCATGACATATTTCCGTACGATTTTCCGATGATTTTACCAATGATCAAGAG 902
Db |||||
QY 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
Db |||||
QY 903 TTTTACTTACCGTCTTAAATAATCGGGAACAAGCTTATAGGATCAATAAATAATCTGGTCTG 962
Db |||||
QY 267 PheThrTyrArgValLysAspArgGluGlnAlaTyrGlyIleAsnLysLysSerGlyLeu 286
Db |||||
QY 963 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGG 1022
Db |||||
QY 287 AsnGluGluIleAsnAsnThrAspLeuIleSerGlnLysTyrTyrIleLeuLysLysGly 306
Db |||||
QY 1023 GAAAGCCGATGATCCCTTTGATCGAGTCACTTGAACCTGTTCCACCATCAATACGTT 1082
Db |||||
QY 307 GluSerProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326
Db |||||
QY 1083 GATGTCGATACCAACGAATTCCTAAAGATGACGACGCTCTTAACAGCTACGGAACGTAAC 1142
Db |||||
```

Db 327 AspValAsnThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
QY 1143 TTAGACTTCAGAGATTATACGATCTCGTGTAAAGCTAAAGTACTCTCAACAATCTC 1202
|||||
Db 347 LeuAspPheArgAspLeuThrAspProCysAspLysAlaLysLeuLeuThrAsnAsnLeu 366
QY 1203 GATGCTTTTGGTATTATGACTATATACCTTTAACTGGAAAAAGTAGAGGATAATCAATCGATGAC 1262
|||||
Db 367 AspAlaPheAspIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspLys 386
QY 1263 ACCAACCGTATCATACCGTTTATATGGCAGACGCCGAGGAGAGAAATGCTAGCTAT 1322
|||||
Db 387 AsnAsnArgIleValThrValTyrMetGlyLysArgProLysGlyAlaLysGlySerTyr 406
QY 1323 CATTTAGCC 1331
|||||
Db 407 HisLeuAla 409

RESULT 8

STRP_STRPY
ID _STRP_STRPY STANDARD; PRT; 440 AA.
AC P10520;
DT 01-JUL-1989 (Rel. 11, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Streptokinase A precursor.
GN Name=ska; OrderedLocusNames=SPY1979;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SF130/13 / Serotype M1;
RX MEDLINE=89160264; PubMed=2646590;
RA Walter F., Siegel M., Malke H.;
RT "Nucleotide sequence of the streptokinase gene from a Streptococcus
pyogenes type 1 strain."
RL Nucleic Acids Res. 17:1261-1261(1989).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=SF370 / ATCC 700294 / Serotype M1;
RX MEDLINE=21192684; PubMed=11296296; DOI=10.1073/pnas.071559398;
RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian X., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.E.;
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
CC -!- FUNCTION: This protein is not a protease, but it activates
plasminogen by complexing with it. As a potential virulence
factor, it is thought to prevent the formation of effective fibrin
barriers around the site of infection, thereby contributing to the
invasiveness of the cells.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
CC -----
CC EMBL; X13399; CA311765.1; -; Genomic DNA.
CC EMBL; AE006620; AAK34665.1; -; Genomic DNA.
CC PIR; S02724; S02724.
CC HSSP; Q53284; 1C4P.
CC SMR; P10520; 63-398.
CC InterPro; IPR004093; Staphylokinase.
CC InterPro; IPR008124; Streptokinase.
CC Pfam; PF02821; Staphylokinase; 3.
CC PRINTS; PR01753; STREPKINASE.
CC Complete proteome; Plasminogen activation; Signal; Virulence.
FT SIGNAL 1 26
FT CHAIN 27 440 Streptokinase A.

FT CONFLICT 163 163 L -> V (in Ref. 1).
FT CONFLICT 345 345 R -> G (in Ref. 1).
FT CONFLICT 373 373 D -> N (in Ref. 1).
FT CONFLICT 428 428 D -> Y (in Ref. 1).
FT CONFLICT 438 438 K -> N (in Ref. 1).
SQ SEQUENCE 440 AA; 49924 MW; D6227BF040B758DB CRC64;

Alignment Scores:
Pred. No.: 3 25e-115 Length: 440
Score: 1794.00 Matches: 341
Percent Similarity: 93.99% Conservative: 19
Best Local Similarity: 89.03% Mismatches: 23
Query Match: 61.04% Indels: 0
DB: 1 Gaps: 0

US-09-940-235-10 (1-1661) x STRP_STRPY (1-440)

QY 183 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACACGCCAATGTGGTGT 242
Db 27 IleAlaGlyTyrGlyTyrLeuProAspArgProIleAsnAsnSerGlnLeuValVal 46
QY 243 AGCGTTGCTGCTACTGTTGAGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC 302
|||||
Db 47 SerMetAlaGlyIleValGluGlyThrAspLysValPheIleAsnPhePheGluIle 66
QY 303 GATCTAACATCACGACCTGCTCATGGAGGAAAGACAGACGAGCTTTAAGTCCAAAATCA 362
|||||
Db 67 AspLeuThrSerGlnProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 86
QY 363 AAACCATTTTGTACTGATGTCGCGGATGTCATATAAACTTGAGAAAGCTGACTTACTA 422
|||||
Db 87 LysProPheAlaThrAspAsnGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu 106
QY 423 AAGGCTATTCAAGAACAAATTGATCGCTTAACAGTCCACAGTAAACGACGACTTTAGGTC 482
|||||
Db 107 LysAlaIleGlnLysGlnLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal 126
QY 483 ATTGATTTCGACAGCGATGCAACATTACTGTCGAACCGGCAAGCTCTACTTTCCTGAC 542
|||||
Db 127 IleAspPheAlaSerAspAlaThrIleThrAspAsnGlyLysValTyrPheAlaAsp 146
QY 543 AAAGATGTTTCGGTAACTTCGCCGACCAACCTGTCCAGAAATTTTGTAAAGCGGACAT 602
Db 147 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuLysGlyHis 166
QY 603 GTCCGCTTAGACCATATAAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 662
Db 167 ValArgValArgProTyrLysGluLysProValGlnAsnGlnAlaLysSerValAspVal 186
QY 663 GANTACTGTACAGTTTACTCCCTTAAACCCGTATGACGATTCAGACCGAGTCTCAA 722
Db 187 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 206
QY 723 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTA 782
Db 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 226
QY 783 GCTCAAGCACAAGCAATTTTAAACAAAAACCCAGGCTATACGATTTATGAACGTCGAC 842
Db 227 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp 246
QY 843 TCCTCAATCGTCACTCATGACATGATTTTCGTCAGATTTTACCAATGGATCAAGAG 902
Db 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
QY 903 TTTTACTTTACCGTGTAAAAAATCGGAAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 962
Db 267 PheThrThrHisValLysAsnArgGluGlnAlaTyrGluLeuAsnProLysThrGlyLe 286
QY 963 AATGAAGAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTCTTAAAAAGGG 1022
Db 287 LysGluLysThrAsnAsnThrAspLeuValSerGluLysTyrTyrValLeuLysGlnGly 306

```
QY 1023 GAAAGCCGTATGATCCCTTTGATCGCAGTCACCTGTGAAACTGTTCCACCATCAATACGTT 1082
Db |||||
QY 307 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326
Db |||||
QY 1083 GATGTCGATACCAACCAATTTGCTAAAGAGTGACGAGCTCTTAACAGCTAGCGAACGTAAC 1142
Db |||||
QY 327 AspValAsnThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
Db |||||
QY 1143 TTAGACTTCAGAGATTATACGATCCCTCGTGATAAGGCTAAACTCTCTCAACAATCTCTC 1202
Db |||||
QY 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
Db |||||
QY 1203 GATGCTTTTGGTATTATGAGCTATACCTTAACCTGGAAGAGTAGAGGATATACGATGAC 1262
Db |||||
QY 367 AspAlaPheAspIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspLys 386
Db |||||
QY 1263 ACCAACCCTATCATACCTTTATATGCGCAAGCGACCCCAAGAGAGAGATGCTAGCTAT 1322
Db |||||
QY 387 AsnAsnArgValValThrValTyrMetGlyLysArgProLysGlyAlaLysGlySerTyr 406
Db |||||
QY 1323 CATTTAGCC 1331
Db |||||
QY 407 HisLeuAla 409
Db |||||
RESULT 9
ID Q53ZX6_STRPY PRELIMINARY; PRT; 440 AA.
AC Q53ZX6.
DT 13-SEP-2005 (T-EMBLrel. 31, Created)
DT 13-SEP-2005 (T-EMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (T-EMBLrel. 31, Last annotation update)
DE Streptokinase.
GN Name=ska;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=86-779.
RX PubMed=14679231; DOI=10.1128/JB.186.1.110-121.2004;
RA Kalia A., Bessen D.E.;
RT "Natural selection and evolution of streptococcal virulence genes
involved in tissue-specific adaptations.";
RL J. Bacteriol. 186:110-121(2004).
DR EMBL; AY234128; AAP39948.1; -; Genomic_DNA.
KW Kinase.
SQ SEQUENCE 440 AA; 49924 MW; D6227EF040B758DB CRC64;
Alignment Scores:
Pred. No.: 3.25e-115 Length: 440
Score: 1794.00 Matches: 341
Percent Similarity: 93.99% Conservative: 19
Best Local Similarity: 89.03% Mismatches: 23
Query Match: 61.04% Indels: 0
DB: 2 Gaps: 0
US-09-940-235-10 (1-1661) x Q53ZX6_STRPY (1-440)
QY 183 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCACCAACAGCCAATTTGGTTGTT 242
Db |||||
QY 27 IleAlaGlyTyrGlyTyrLeuProAspArgProPheIleAsnAsnSerGlnLeuValVal 46
Db |||||
QY 243 AGCGTGTCTGGTACTGTTGAGGGGAGCAATCAAGACATAGTCTTAAATTTTTTGAATC 302
Db |||||
QY 47 SerMetAlaGlyIleValGluGlyThrAspLysLysValPheIleAsnPheGluIle 66
Db |||||
QY 303 GATCTAAACATCAACGCTCTCATGAGGAGAAAGACAGACGAGCTTAAGTCCAAATCA 362
Db |||||
QY 67 AspLeuThrSerGlnProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 86
Db |||||
QY 363 AAACCATTTGCTACTGATAGTGGCGGATGTCACATAAACTTGAGAAAGCTGACTTACTA 422
Db |||||
```

```
Db 87 LysProPheAlaThrAspAsnGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu 106
QY 423 AAGCTATTCAAGAACAACTTCATCGTAACTGTCACACTACGACGACTACTTTGAGTGC 482
Db |||||
QY 107 LysAlaIleGlnLysGlnLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal 126
Db |||||
QY 483 ATTGATTTTGCACGATGCAACCATCTACTGATCGAAACGGCAAGTCTACTTTGCTGAC 542
Db |||||
QY 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146
Db |||||
QY 543 AAAAGATGTTCCGTAACCTTCCCGACCCCACTGCTCAAGATTTTCTAAGCGGACAT 602
Db |||||
QY 147 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLysGlyHis 166
Db |||||
QY 603 GTGCGCTGTAGACCATATAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 662
Db |||||
QY 167 ValArgValArgProTyrLysGlyLysProValGlnAsnGlnAlaLysSerValAspVal 186
Db |||||
QY 663 GAATATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTCAGACCAAGTCTCAAA 722
Db |||||
QY 187 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 206
Db |||||
QY 723 GATTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAACTACTA 782
Db |||||
QY 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 226
Db |||||
QY 783 GCTCAAGCACAAGCATTTTAAACAAACCAACCCAGCTATACGATTATGACGTCAC 842
Db |||||
QY 227 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp 246
Db |||||
QY 843 TCCTCAATCGTCACTCATGATGACATGACATTTTCGTCGATGTTTACCAATCGATCAAGAG 902
Db |||||
QY 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuPhePheArgGlnGlu 266
Db |||||
QY 903 TTTACTTACCGTGTAAAAATTCGGGAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 962
Db |||||
QY 267 PheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsnProLysThrGlyIle 286
Db |||||
QY 963 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAGTCTTAAAAAAGG 1022
Db |||||
QY 287 LysGluLysThrAsnAsnThrAspLeuValSerGluLysTyrTyrValLeuLysGlnGly 306
Db |||||
QY 1023 GAAAGCCGTATGATCCCTTTGATCGCAGTCACCTGTGAAACTGTTCCACCATCAATACGTT 1082
Db |||||
QY 307 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326
Db |||||
QY 1083 GATGTCGATACCAACCAATTTGCTAAAGAGTGACGAGCTCTTAAACAGCTAGCGAACGTAAC 1142
Db |||||
QY 327 AspValAsnThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
Db |||||
QY 1143 TTAGACTTCAGAGATTATACGATCCCTCGTGATAAGGCTAAACTCTCTCAACAATCTCTC 1202
Db |||||
QY 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
Db |||||
QY 1203 GATGCTTTTGGTATTATGAGCTATACCTTAACCTGGAAGAGTAGAGGATATACGATGAC 1262
Db |||||
QY 367 AspAlaPheAspIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspLys 386
Db |||||
QY 1263 ACCAACCCTATCAACCGTTTATATGCGCAAGCGACCCCAAGAGAGAGATGCTAGCTAT 1322
Db |||||
QY 387 AsnAsnArgValValThrValTyrMetGlyLysArgProLysGlyAlaLysGlySerTyr 406
Db |||||
QY 1323 CATTTAGCC 1331
Db |||||
QY 407 HisLeuAla 409
Db |||||
RESULT 10
Q7X0Y2_STRPY
ID Q7X0Y2_STRPY PRELIMINARY; PRT; 440 AA.
AC Q7X0Y2.
DT 01-OCT-2003 (T-EMBLrel. 25, Created)
DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
```

Qy	723	GATCTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAATTACTA	783
Db	207	AspThrLysLeuLeuLysLysLeuAlaIleGlyAspThrValThrSerGlnGluLeuLeu	226
Qy	783	GCTCAACGACAAAGCATTTTAAACAAAAACCACCGCTATATCGATTATGAACGCTGAC	842
Db	227	AlaGlnAlaGlnSerIleLeuAsnGluSerHisProAspTyrThrIleTyrGluArgAsp	246
Qy	843	TCCTCAATCGTCACTCATATGCAATTTTCCGTACGATTTTACCAATGGATCAAGAG	902
Db	247	SerSerileValThrHisAspLysAspIlePheargThrIleLeuProMetAspGlnGlu	266
Qy	903	TTTACTTACCGTGTATAAAATCGGGAAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG	962
Db	267	PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu	286
Qy	963	AATGAAGAATATAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG	1022
Db	287	AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysArg	306
Qy	1023	GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACATGTTCCACCATCAAAATAGCTT	1082
Db	307	GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal	326
Qy	1083	GATGTCGATACCAACGAATTCGTAAAGAGTGCAGCAGCTCTTAAACAGCTAGCGAACGTAAC	1142
Db	327	AspValAsnThrAsnGluLeuLeuLysLysSerGluGlnLeuLeuThrAlaSerGluArgAsn	346
Qy	1143	TTAGACTTTCAGAGATTTATAGGATCCTCGTGATAGAGGCTAAACTACTCTACACAATCTC	1202
Db	347	LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu	366
Qy	1203	GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAAGTAGAGGATAATCAGCATGAC	1262
Db	367	AspAlaPheAspIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspLys	386
Qy	1263	ACCAACCGTATCATCAACCGTTTATATGGCAAGCGACCCGAAGAGAGAGAATGCTAGCTAT	1322
Db	387	AsnAsnArgValValThrValTyrMetGlyLysArgProLysGlyAlaLysGlySerTyr	406
Qy	1323	CATTTAGCC 1331	
Db	407	HisLeuAla 409	
RESULT 11			
Q7X0Y7_STRPY PRELIMINARY; PRT; 440 AA.			
AC	Q7X0Y7;		
DT	01-OCT-2003 (TrEMBLrel. 25, Created)		
DT	01-OCT-2003 (TrEMBLrel. 25, Last sequence update)		
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)		
DE	Streptokinase.		
GN	Name=Ska;		
OS	Streptococcus pyogenes.		
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;		
OC	Streptococcus.		
OX	NCBI_TaxID=1314;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RC	STRAIN=1RP112;		
RC	PubMed=14679231; DOI=10.1128/JB.186.1.110-121.2004;		
RA	Kalia A., Bessen D.E.;		
RT	Natural selection and evolution of streptococcal virulence genes		
RT	involved in tissue-specific adaptations.;		
RL	J. Bacteriol. 186:110-121(2004).		
RL	EMBL; AY234130; AAP39950.1; -; Genomic_DNA.		
DR	HSP; Q53284; 1C4P.		
DR	SMR; Q7X0Y7; 63-398.		
DR	GO; GO:0016301; P:kinase activity; IEA.		
DR	GO; GO:0008243; F:plasmogen activator activity; IEA.		
DR	GO; GO:0005515; F:protein binding; IEA.		
DR	InterPro; IPR004093; Staphylokinase.		
DR	InterPro; IPR008124; Streptokinase.		

US-09-940-235-10 (1-1661) x Q7X0Y2_STRPY (1-440)

QY	183	ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAAATGGTTGTT	242
Db	27	IleAlaGlyTyrGlyTyrLeuProAspArgProIleAsnAsnSerGlnLeuValVal	46
QY	243	AGCGTCTGCTACTGTTTGGGGGACGAATCAAGACATTAGTCTTAAATTTTGTGAATTC	302
Db	47	SerMetAlaGlyIleValGlnGlyThrAspLysLysValPheIleAsnPheGluIle	66
QY	303	GATCTTAACATCACGACCTGCTCATCGAGGAAAGACAGACGACGAGGCTTAAGTCCAAATCA	362
Db	67	AspLeuThrSerGlnProAlaHisGlyGlyThrGlnGlnGlyLeuSerProIysSer	86
QY	363	AAACCATTTGCTACTGATAGTGGCGCGATGTCAACATAACTTTGAGAAACCTGACTTACTTA	422
Db	87	LysProPheAlaThrAspAsnSerAlaMetProHisLysLysLeuGluLysAlaAspLeuLeu	106
QY	423	AAGGCTATTCAAGAACAAATTGATTCGCTAAAGTCCACAGTAACGACGACTACTTTGAGGTC	482
Db	107	LysAlaIleGlnLysGlnLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal	126
QY	483	ATTGATTTTGCACGGATGCACCACTTACTGATCGAAACGGCAAGTCTACTTTGCTGCAC	542
Db	127	IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp	146
QY	543	AAAGATGGTTCGGTAACCTTGCCCGACCAACCTGTCCAGAAATTTTGTCTAAGCGGACAT	602
Db	147	LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuThrGlyHis	166
QY	603	GTGCGCGTTAGACCATATAAAGAAAAACAATATACAAAACCAAGCGGAAATCTCTGTGATGTG	662
Db	167	ValArgValArgProTyrLysGluLysProValGlnAsnGlnAlaLysSerValAspVal	186
QY	663	GAATATACTGTACAGTTTACTCCCTTAAACCCGTGATGAGATTTCAGACCAAGGTCCTCAA	722
Db	187	GluTyrThrValGlnPheThrProLeuAsnProAspAspAspPheLysProValLeuLys	206

DR Pfam: PF02821; Staphylokinase; 3.
DR PRINTS; PR01753; STREPKINASE.
KW Kinase.
SQ SEQUENCE 440 AA; 50160 MW; 1C5AF07907EC7AC8 CRC64;

Alignment Scores:
Pred. No.: 9,86e-113 Length: 440
Score: 1758.00 Matches: 338
Percent Similarity: 93.21% Conservatives: 19
Best Local Similarity: 88.25% Mismatches: 26
Query Match: 59.82% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-10 (1-1661) x QYX0Y7_STRPY (1-440)

QY 183 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCCAATTTGTTGTT 242
DB 27 IleAlaGlyTyrGlyTrpLeuLeuAspArgProValAsnAsnSerGlnLeuVal 46

QY 243 AGCGTTGCTGTTACTGTTAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC 302
DB 47 SerMetAlaGlyIleValGluGlyThrAspLysLysValPheIleAsnPheGluIle 66

QY 303 GATCTAACATCAGCACTGCTCATGAGGAGAAACAGACAGCAAGCTTAAAGTCCAAATCA 362
DB 67 AspLeuThrSerGlnProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 86

QY 363 AAACCAATTTGCTACTGATGTCGGGATCTCATATAAATCTTGAGAAGCTGACTTACTA 422
DB 87 LysProPheAlaThrAsnSerSerAlaMetProHisLysLeuGluLysAlaAspLeuLeu 106

QY 423 AAGGCTATTCAAGAACAAATTTGATCGCTTAACGTCACAGTACACGACGACTACTTTGAGTC 482
DB 107 LysAlaIleGlnGluArgLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal 126

QY 483 ATTGATTTTGAGCGATGCAACCATTTACTGATCGGAAACCGCAAGTCTACTTTGTCGAC 542
DB 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146

QY 543 AAAGATGGTTCGTACCTTCCGACCCCACTGTCACAGATTTTTCGCTTAAAGCGGCAT 602
DB 147 LysAspAspSerValThrLeuProThrGlnProValGlnGluPheLeuLeuArgGlyHis 166

QY 603 GTGCGGTTAGACCATATATAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 662
DB 167 ValArgValArgProTyrLysGluLysProIleGlnThrProAlaLysSerValAspVal 186

QY 663 GAATATACGTACAGTTTACTCCCTTAAACCTGATGACGATTTTCAGACAGGTCTCTCAA 722
DB 187 ArgTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProValLeuLys 206

QY 723 GATACCTAAGCTATTGAAACACTAGTATCGGTGACACCATCATCTCAAGAAATTTACTA 782
DB 207 AsnThrLysLeuLeuLysThrLeuAlaIleGlyGlyThrValThrSerGlnGluLeuLeu 226

QY 783 GCTCAAGCAAGACATTTTAAACAAAACCAACCCAGGCTATACGATTTTATGACGTGAC 842
DB 227 AlaGlnAlaGlnSerIleLeuAsnGlnSerHisProAspTyrThrIleTyrGluArgAsp 246

QY 843 TCCTCAATCGTCACTCATGCAATGCAATTTTCCGTACGATTTTACCAATGGATCAAGAG 902
DB 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266

QY 903 TTTACTACCGTGTAAATCGGGAACAGCTTATAGGATCAATAAAAAATCTGCTGTCG 962
DB 267 PheThrTyrHisIleLysAspArgGluGlnAlaTyrGlyIleAsnLysLysSerGlyGln 286

QY 963 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTTACGCTCTTAAAAAGGG 1022
DB 287 GluGluLysThrAsnAsnThrAspLeuIleSerGluLysTyrThrValLeuLysLysGly 306

QY 1023 GAAAAGCCGTATGATCCCTTTTCATCGCAGTCACTTCAAACTGTTTCCACCATCAATACGTT 1082
DB 243 AGCGTTGCTGTTACTGTTAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC 302

307 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleAsnTyrVal 326
1083 GATGTCGATACCAACGAATTTGCTTAAAGTGGACAGCTCTTAAACAGCTAGCGAAGTAAAC 1142
DB 327 AspValAsnThrAsnLysLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
1143 TTAGACTTCAGAGATTTTATACGATCCTCGTGATAGGCTAAACTACTCTACAAACAATCTC 1202
DB 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
1203 GATGCTTTTGTATATGACTATACCTTAACCTGGAAGAGTAGAGGATAATCAGCATGAC 1262
DB 367 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspLys 386
1263 ACCAACGCTATCATACCGTTTATATGCGCAAGCGACCGGAGGAGAGAGAACTAGCTAT 1322
DB 387 AsnAsnArgValValThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 406
1323 CATTTAGCC 1331
DB 407 HisLeuAla 409

RESULT 12
Q5X9T6_STRP6
ID Q5X9T6_STRP6 PRELIMINARY; PRT; 440 AA.
AC Q5X9T6;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Streptokinase (SC 3.4.5.-).
GN OrderedLocustNames=M6_Spy1692;
OS Streptococcus pyogenes (serotype M6).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=301450;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MGAS10394;
RX PubMed=15272401; DOI=10.1086/422697;
RA Banks D.J., Porcella S.F., Barbican K.D., Beres S.B., Phillips L.E.,
RA Voyich J.M., Deleo F.R., Martin J.M., Somerville G.A., Musser J.M.;
RT "Progress toward characterization of the group A Streptococcus
metagenome: complete genome sequence of a macrolide-resistant serotype
M6 strain."
RL J. Infect. Dis. 190:727-738 (2004).
DR EMBL; CP000003; AAT87827.1; -; Genomic_DNA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008243; F:plasmaenogen activator activity; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR InterPro; IPR004093; Staphylokinase.
DR InterPro; IPR008124; Streptokinase.
DR Pfam; PF02821; Staphylokinase; 3.
DR PRINTS; PR01753; STREPKINASE.
KW Complete proteome; Hydrolase; Kinase.
SQ SEQUENCE 440 AA; 50160 MW; 1C5AF07907EC7AC8 CRC64;

Alignment Scores:
Pred. No.: 9,86e-113 Length: 440
Score: 1758.00 Matches: 338
Percent Similarity: 93.21% Conservatives: 19
Best Local Similarity: 88.25% Mismatches: 26
Query Match: 59.82% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-10 (1-1661) x Q5X9T6_STRP6 (1-440)

QY 183 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCCAATTTGTTGTT 242
DB 27 IleAlaGlyTyrGlyTrpLeuLeuAspArgProValAsnAsnSerGlnLeuVal 46

QY 243 AGCGTTGCTGTTACTGTTAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC 302
DB 47 SerMetAlaGlyIleValGluGlyThrAspLysLysValPheIleAsnPheGluIle 66

QY 303 GATCTAACATCAGCACTGCTCATGAGGAGAAACAGACAGCAAGCTTAAAGTCCAAATCA 362
DB 67 AspLeuThrSerGlnProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 86

QY 363 AAACCAATTTGCTACTGATGTCGGGATCTCATATAAATCTTGAGAAGCTGACTTACTA 422
DB 87 LysProPheAlaThrAsnSerSerAlaMetProHisLysLeuGluLysAlaAspLeuLeu 106

QY 423 AAGGCTATTCAAGAACAAATTTGATCGCTTAACGTCACAGTACACGACGACTACTTTGAGTC 482
DB 107 LysAlaIleGlnGluArgLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal 126

QY 483 ATTGATTTTGAGCGATGCAACCATTTACTGATCGGAAACCGCAAGTCTACTTTGTCGAC 542
DB 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146

QY 543 AAAGATGGTTCGTACCTTCCGACCCCACTGTCACAGATTTTTCGCTTAAAGCGGCAT 602
DB 147 LysAspAspSerValThrLeuProThrGlnProValGlnGluPheLeuLeuArgGlyHis 166

QY 603 GTGCGGTTAGACCATATATAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 662
DB 167 ValArgValArgProTyrLysGluLysProIleGlnThrProAlaLysSerValAspVal 186

QY 663 GAATATACGTACAGTTTACTCCCTTAAACCTGATGACGATTTTTCAGACAGGTCTCTCAA 722
DB 187 ArgTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProValLeuLys 206

QY 723 GATACCTAAGCTATTGAAACACTAGTATCGGTGACACCATCATCTCAAGAAATTTACTA 782
DB 207 AsnThrLysLeuLeuLysThrLeuAlaIleGlyGlyThrValThrSerGlnGluLeuLeu 226

QY 783 GCTCAAGCAAGACATTTTAAACAAAACCAACCCAGGCTATACGATTTTATGACGTGAC 842
DB 227 AlaGlnAlaGlnSerIleLeuAsnGlnSerHisProAspTyrThrIleTyrGluArgAsp 246

QY 843 TCCTCAATCGTCACTCATGCAATGCAATTTTCCGTACGATTTTACCAATGGATCAAGAG 902
DB 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266

QY 903 TTTACTACCGTGTAAATCGGGAACAGCTTATAGGATCAATAAAAAATCTGCTGTCG 962
DB 267 PheThrTyrHisIleLysAspArgGluGlnAlaTyrGlyIleAsnLysLysSerGlyGln 286

QY 963 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTTACGCTCTTAAAAAGGG 1022
DB 287 GluGluLysThrAsnAsnThrAspLeuIleSerGluLysTyrThrValLeuLysLysGly 306

QY 1023 GAAAAGCCGTATGATCCCTTTTCATCGCAGTCACTTCAAACTGTTTCCACCATCAATACGTT 1082
DB 243 AGCGTTGCTGTTACTGTTAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC 302


```

Db 47 SerMetAlaGlyIleValGluGlyThrAspLysLysValPheIleAsnPhePheGluIle 66
QY 303 GATCTAAACATCAGCAGCTCTCATGAGGAGAAACAGACAGAGGCTTAAGTCCAAATCA 362
Db 67 AspLeuThrSerGlnProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 86
QY 363 AAACCATTTGCTACTAGTGGCGGATGTGCATATAAACTTCAGAGAAAGCTGACTACTA 422
Db 87 LysProPheAlaThrAsnSerSerAlaMetProHisLysLeuGluLysAlaAspLeuLeu 106
QY 423 AAGGCTATTCAAGAACAAATGATCGCTAACGCTCCACAGTAACGACGACTACTTTGAGTTC 482
Db 107 LysAlaIleGlnGluArgLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal 126
QY 483 ATTGATTTTGCAGGATGCACCATTTACTGATCGAAGCGCAAGTCTACTTTGCTGAC 542
Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146
QY 543 AAAGATGGTTTCGGTAACCTTTCGCGACCCCACTCTGCCAAGATTTTTCGTAAGCGGACAT 602
Db 147 LysAspAspSerValThrLeuProThrGlnProValGlnGluPheLeuLeuArgGlyHis 166
QY 603 CTGCGCGTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 662
Db 167 ValArgValArgProTyrLysGluLysProIleGlnThrProAlaLysSerValAspVal 186
QY 663 GAATATCTGATGACTTTACTCCCTTAACCTTCATGATGACGATTTGACAGGCTCTCAA 722
Db 187 ArgTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProValLeuLys 206
QY 723 GATCTAAGCTATTGAAACACTAGCTATCGTGTGACCATCACATCTCAAGAAATCTACTA 782
Db 207 AsnThrLysLeuLeuLysThrLeuAlaIleGlyGlyThrValThrSerGlnGluLeuLeu 226
QY 783 GCTCAAGCACAAAGCATTTTAAACAAAACCAACCCAGGCTATACGATTTTATGAACCTGAC 842
Db 227 AlaGlnAlaGlnSerIleLeuAsnGluSerHisProAspTyrThrIleTyrGluArgAsp 246
QY 843 TCCTCAATCGTCATCATGACAATGACATTTCCGTACGATTTTACCAATGGAATCAAGAG 902
Db 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
QY 903 TTTACTTACCGTGTAAATAACCGGAACAGCTTATAGGATCAATAAATAAATCTGCTCTG 962
Db 267 PheThrTyrHisIleLysAspArgGluGlnAlaTyrGlyIleAsnLysLysSerGlyGln 286
QY 963 AATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGCTCTTAAAAAGGG 1022
Db 287 GluGluLysThrAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 306
QY 1023 GAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAACACTGTTCCACCAATAATACGTT 1082
Db 307 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleAsnTyrVal 326
QY 1083 GATCTCGATACCAACGAATGCTAAAGTAGGAGCAGCTCTTAAACAGCTAGCGAACCTTAC 1142
Db 327 AspValAsnThrAsnLysLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
QY 1143 TTAGACTTTCAGAGATTATACCATCTCTGATGATAGGCTTAACTACTCTACCAACATCTC 1202
Db 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
QY 1203 GATGCTTTTGGTATTATGACTATATACCTTAACCTGGAAGAGTAGAGGATATCACGATGAC 1262
Db 367 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspLys 386
QY 1263 ACCAACCGTATCATACCGTTTATATGGCAAGCGCCGAGGAGAGAGATGCTAGCTAT 1322
Db 387 AsnAsnArgValValThrValTyrMetGlyLysArgProGluGluAsnAlaSerTyr 406
QY 1323 CATTTAGCC 1331

```

```

Db 407 HisLeuAla 409
RESULT 13
Q8NZA6_STRP8
ID Q8NZA6_STRP8 PRELIMINARY; PRT; 440 AA.
AC Q8NZA6;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Streptokinase.
GN OrderedLocustNames-spyM18_2042;
OS Streptococcus pyogenes (serotype M18).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=301451;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MGAS232 / Serotype M18;
RX MEDLINE=21927593; PubMed=11917108; DOI=10.1073/pnas.062526099;
RA Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
RA Sylva G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F., Zhang Q.,
RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
RT "Genome sequence and comparative microarray analysis of serotype M18
RT group A Streptococcus strains associated with acute rheumatic fever
outbreaks.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).
DR EMBL; AS010107; ALU98517.1; -; Genomic_DNA.
DR HSSP; Q53284; 1C4P.
DR SNR; Q8NZA6; 63-398.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008243; F:plasma membrane activator activity; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0004093; Streptokinase.
DR InterPro; IPR008124; Streptokinase.
DR Pfam; PF02821; Staphylokinase; 3.
DR PRINTS; PR01753; STREPKINASE.
DR Complete proteome; Kinase.
KW SEQUENCE 440 AA; 50382 MW; 47DAB35577EBB88E CRC64;
SQ
Alignment Scores:
Pred. No.: 1,596-112 Length: 440
Score: 1755.00 Matches: 334
Percent Similarity: 93.73% Conservative: 25
Best Local Similarity: 87.21% Mismatches: 24
Query Match: 59.71% Indels: 0
DB: Gaps: 2
US-09-940-235-10 (1-1661) x Q8NZA6_STRP8 (1-440)
QY 183 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAAATGCTGTT 242
Db 27 IleAlaGlyTyrGluTyrLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 46
QY 243 AGGTTGCTGGTACTCTGTGAGGGGAGCAATCAAGACATAGTCTTAAATTTTGAATC 302
Db 47 SerMetAlaGlyIleValGluGlyThrAspLysValPheIleAsnPheGluIle 66
QY 303 GATCTAACATCAACGACTCTCATGAGGAGAAACAGACAGAGGCTTAAGTCCAAATCA 362
Db 67 AspLeuThrSerGlnProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 86
QY 363 AAACCATTTGCTACTAGTGGCGGATGTGCATATAAACTTCAGAGAAAGCTGACTACTA 422
Db 87 LysProPheAlaThrAspAsnGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu 106
QY 423 AAGGCTATTCAAGAACAAATTTGATCGCTAACGCTCCACAGTAACGACGACTACTTTGAGTTC 482
Db 107 LysAlaIleGlnGluArgLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal 126
QY 483 ATTGATTTTGCAGGATGCACCATTTACTGATCGAAGCGCAAGTCTACTTTGCTGAC 542
Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsn 146

```


QY 543 AAGATGTTGGTAACTTGGCCGACCAACCTGTCTCAAGAAATTTTGTCTAAGCGGACAT 602
 Db :::
 147 GlnAspGlySerValThrLeuProThrGlnProIleGlnPheLeuLeuArgGlyHis 166
 QY 603 GTGCGGTAGACCATATAAGAAACCAATACAAACCAACGGAATCTGTGTGATGTG 662
 Db :::
 167 ValArgValArgProTyrLysGluLysProIleGlnThrProAlaLysSerValAsp 186
 QY 663 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATCAGCATTTTCAGACGAGTCTCAA 722
 Db :::
 187 ArgTyrThrValGlnPheThrProLeuAsnProAspAspPheLysProValLeuLys 206
 QY 723 GATCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTA 782
 Db :::
 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 226
 QY 783 GCTCAAGCACAAGCAATTTTAAACCAAAACCAACCCAGGCTATACGATTTATGAACGTGAC 842
 Db :::
 227 AlaGlnAlaGlnSerIleLeuAsnGluSerHisSerAspTyrThrIleTyrGluArgAsp 246
 QY 843 TCCTCAATCGTCTCATGACATGACATTTTCGTACGATTTTACCAATGGATCAAGAG 902
 Db :::
 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
 QY 903 TTTACTTACCGTGTAAAAATCGGGAACCAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 962
 Db :::
 267 PheThrTyrHisIleLysAspArgGluGlnAlaTyrGlyIleAsnLysLysSerGlyGln 286
 QY 963 AATGAAGAATAAACAACCTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1022
 Db :::
 287 GluGluLysThrAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 306
 QY 1023 GAAAGCCGTATGATCCCTTGTGTCGAGCTCACTTGAACCTGTTCAACATCAATAGTT 1082
 Db :::
 307 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleAsnTyrVal 326
 QY 1083 GATGTCGATACCAACCAATTTGCTTAAAGAGTACGAGCTCTTAAACAGCTAGCAACGTAAC 1142
 Db :::
 327 AspValAsnThrAsnLysLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
 QY 1143 TTAGACTTCAGAGATTTATACGATCTCTGATGATGAGCTTAACTCTTCAACAATCTC 1202
 Db :::
 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
 QY 1203 GATGCTTTGGTATTATGACTATACCTTAACTGGAAGTAGAGGATATACGATGAC 1262
 Db :::
 367 AspAlaPheGlyLeuMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspLys 386
 QY 1263 ACCAACCGTATCATACCGTTTATATGCGCAAGCGACCGAGAGAGATGCTAGCTAT 1322
 Db :::
 387 AsnAsnArgValThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 406
 QY 1323 CATTTAGCC 1331
 Db :::
 407 HisLeuAla 409

RESULT 14
 Q7X0Y3_STRPY PRELIMINARY; PRT; 440 AA.
 AC Q7X0Y3.
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Streptokinase.
 GN Name=ska;
 OS Streptococcus pyogenes.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1314;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=ALAB49;

RX PubMed=14679231; DOI=10.1128/JB.186.1.110-121.2004;
 RA Kalra A., Bessen D.E.;
 RT "Natural selection and evolution of streptococcal virulence genes
 involved in tissue-specific adaptations.";
 RL J. Bacteriol. 186:110-121(2004).
 DR EMBL; AY234134; AAPJ9954.1; -; Genomic_DNA.
 DR HSSP; Q53284; 1C4P.
 DR SMR; Q7X0Y3; 63-398.
 DR GO; GO:0016301; F:kinase activity; IEA.
 DR GO; GO:0008243; F:plasminogen activator activity; IEA.
 DR GO; GO:0005545; F:protein binding; IEA.
 DR InterPro; IPR004093; Staphylokinase.
 DR InterPro; IPR008124; Streptokinase.
 DR Pfam; PF02821; Staphylokinase; 3.
 DR PRINTS; PR01753; STREPKINASE.
 KW Kinase.
 SQ SEQUENCE 440 AA; 50342 MW; B3C60BAA50DB2C39 CRC64;

Alignment Scores:
 Pred. No.: 2.55e-110 Length: 440
 Score: 1723.00 Matches: 327
 Percent Similarity: 92.43% Conservatives: 27
 Best Local Similarity: 85.38% Mismatches: 29
 Query Match: 58.63% Indels: 0
 DB: 2 Gaps: 0

US-09-940-235-10 (1-1661) x Q7X0Y3_STRPY (1-440)

QY 183 ATTCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACCAACGCAATTTGTTT 242
 Db 27 IleAlaGlyTyrGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 46
 QY 243 AGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 302
 Db 47 SerMetAlaGlyIleValGluGlyThrAspLysLysValPheIleAsnPheGluIle 66
 QY 303 GATCTAACATCACCACCTGCTCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 362
 Db 67 AspLeuThrSerGlnProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLys 86
 QY 363 AAACATTTGCTACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 422
 Db 87 LysProPheAlaThrAspAsnGlyAlaMetProHisLysLeuGluLysAlaAspLeu 106
 QY 423 AAGCTATTCAAGAACAAATGATCGCTAACGTCACAGTAACGACGACTCTTTGAGTGC 482
 Db 107 LysAlaIleGlnGluArgLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal 126
 QY 483 ATTGATTTTCCAGCGATGCAACCATTTACTGATCGAAACCGCAAGGCTCTACTTTGCTGAC 542
 Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAspAsnIleTyrPheAlaAsn 146
 QY 543 AAAGATGTTTCGGTAACTTTGCCGACCAACCTGTCCAAGAAATTTTGTAAAGCGGACAT 602
 Db 147 GlnAspGlySerValThrLeuProThrGlnProIleGlnPheLeuLeuArgGlyHis 166
 QY 603 GTGCGCTTTAGACCATATAAGAAAAACCAATACAAACCAACGCGAATCTGTGTGATGTG 662
 Db 167 ValArgValArgProTyrLysGluLysProIleGlnThrProAlaLysSerValAsp 186
 QY 663 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATCAGCATTTTCAGACCGGCTCAA 722
 Db 187 ArgTyrThrValGlnPheThrProLeuAsnProAspAspPheLysProValLeuLys 206
 QY 723 GATCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTA 782
 Db 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 226
 QY 783 GCTCAAGCACAAGCAATTTTAAACCAAAACCAACCCAGGCTATACGATTTATGAACGTGAC 842
 Db 227 AlaGlnAlaGlnSerIleLeuAsnGluSerHisProAspTyrThrIleTyrGluArgAsp 246
 QY 843 TCCTCAATCGTCTCATGACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 902

Db 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProThrAspGlnGlu 266
 QY 903 TTTACTACCGTGTAAATAATCGGAAACAAGCTTATAGATCAATAAAAAATCTGGTCTG 962
 Db 267 PheThrTyHisValLysAsnArgGluGlnAlaTyrglnAsnAspAsnLysThrGlyLeu 286
 QY 963 AATGAAGAAATAACAACACTGACCTGATCTCTGAGAAATATTACCTCTTAAAGAGG 1022
 Db 287 LysLysGluThrLysAsnThrAspLeuIleSerGluLysTyrrIleLeuLysLysGly 306
 QY 1023 GAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTCAAACTGTTCACCAATCAATACGTT 1082
 Db 307 GluLysProTyAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrrVal 326
 QY 1083 GATGTCGATACCAACGAATTGTCTGAGAGAGCTCTTAAACAGCTAGCGAACGTAAC 1142
 Db 327 AspValAspThrLysAspLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
 QY 1143 TTAGACTTCAGAGATTATACGATCTCTGATAGGCTTAAGCTTAACACTCTTCAACAATCTC 1202
 Db 347 LeuAspPheArgAspLeuTyrrAspProArgAspLysAlaLysLeuLeuTyrrAsnAsnLeu 366
 QY 1203 GATGCTTTTGGTATTATGAGTATACCTTAACTTGGAAAGTAGAGGATAATCACGATGAC 1262
 Db 367 AspAlaPheAspIleMetAspTyrrThrLeuThrGlyLysValGluAspAsnHisAspLys 386
 QY 1263 ACCAAGCTGATACATACCGTTTATATGGCAAGCGACCGAAGAGGAGAGAAATGCTAGTAT 1322
 Db 387 AsnAsnArgValValThrValTyrrMetGlyLysArgProLysGlyAlaLysGlySerTyrr 406
 QY 1323 CATTTAGCC 1331
 Db 407 HisLeuAla 409

RESULT 15
 Q7X0Y0_STRPY PRELIMINARY; PRT; 440 AA.
 AC Q7X0Y0;
 DT 01-OCT-2003 (TremBLrel. 25, Created)
 DT 01-OCT-2003 (TremBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TremBLrel. 26, Last annotation update)
 DE Streptokinase.
 GN Name=ska;
 OS Streptococcus pyogenes.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1314;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=d633;
 RX PubMed=14679231; DOI=10.1128/JB.186.1.110-121.2004;
 RA Kalia A., Bessen D.E.;
 RT "Natural selection and evolution of streptococcal virulence genes involved in tissue-specific adaptations.";
 RL J. Bacteriol. 186:110-121(2004).
 DR EMBL; AY234139; AAP39958.1; -; Genomic_DNA.
 DR HSSP; Q53284; 1C4P.
 DR SMR; O7X0Y0; 63-398.
 DR GO; GO:0016301; F:kinase activity; IEA.
 DR GO; GO:0008243; F:plasminogen activator activity; IEA.
 DR GO; GO:0005515; F:protein binding; IEA.
 DR InterPro; IPR004093; Staphylokinase.
 DR InterPro; IPR008124; Streptokinase.
 DR Pfam; PF02821; Staphylokinase; 3.
 DR PRINTS; PR01753; STREPKINASE.
 KW Kinase.
 SQ SEQUENCE 440 AA; 50143 MW; 4ACB2C29349D680C CRC64;

Alignment Scores:
 Pred. No.: 9, 09e-110 Length: 440
 Score: 1715.00 Matches: 325
 Percent Similarity: 91.91% Conservative: 27

Best Local Similarity: 84.86% Mismatches: 31
 Query Match: 58.35% Indels: 0
 DB: 2 Gaps: 0
 US-09-940-235-10 (1-1661) x Q7X0Y0_STRPY (1-440)
 QY 183 ATTGCTGGACCTCAGTGGCTGTAGACCGTCCATCTGTCAACACACCAATGGTGT 242
 Db 27 IleAlaGlyProGluThrLeuLeuGlyArgProProValAsnAsnSerGlnLeuVal 46
 QY 243 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAATC 302
 Db 47 SerMetAlaGlyIleValGluGlyThrAspLysLysValPheIleAsnPhePheGluIle 66
 QY 303 GATCTAACATCAGACCTGCTCATGAGGAAACAGACAGCAAGCTTAAAGTCCAAATCA 362
 Db 67 AspLeuThrSerGlnProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 86
 QY 363 AAACCATTTGCTACTGATGCGCGATGCTCATATAAACTTGAGAAAGCTGACTTACTA 422
 Db 87 LysProPheAlaThrAspAsnGlyAlaMetProHisLysLeuGluLysAlaAspLeu 106
 QY 423 AAGGCTATTCAAGAAACAATTCATCGCTTAACGTCACAGTAACGACCACTACTTTGAGGTC 482
 Db 107 LysAlaIleGlnGluArgLeuIleAlaAsnValHisSerAsnAspGlyTyrrPheGluVal 126
 QY 483 ATTGATTTTCAAGCGATGCAACATTACTGATCGAAACGGCAAGGTCCTACTTTGCTGAC 542
 Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAspAspAsnIleTyrrPheAlaAsn 146
 QY 543 AAAGATGTTTCGGTAAACCTTGCAGCCCACTGTCACAGAAATTTTGTAAAGCGGACAT 602
 Db 147 GlnAspGlySerValThrLeuProThrGlnProIleGlnPheLeuLeuArgGlyHis 166
 QY 603 GTGCGGTTAGACCATATAAGAAACCAATAACAAACCAAGCGAAATCTGTGATGTG 662
 Db 167 ValArgValArgProTyrrLysGluLysProIleGlnThrProAlaLysSerValAspIle 186
 QY 663 GAATATCTGTACAGTTTACTCCCTTAAACCTTGATGACGATTCAGACAGGTCCTCAA 722
 Db 187 ArgTyrrThrValGlnPheThrProLeuAsnProAspAspAspPheLysProValLeuLys 206
 QY 723 GATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCATCTCAAGATTTACTA 782
 Db 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAsnThrIleThrSerGlnGluLeu 226
 QY 783 GCTCAAGCACAAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTATGAACGTGAC 842
 Db 227 AlaGlnAlaGlnSerIleLeuAsnGluSerHisProAspTyrrThrIleTyrrGluArgAsp 246
 QY 843 TCCTCAATCTGCTCATCATGCAATTCATTTCCGTACGATTTTACCAATGGATCAAGAG 902
 Db 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProThrAspGlnGlu 266
 QY 903 TTTACTACCGTGTAAATAATCGGAAACAAGCTTATAGATCAATAAAAAATCTGGTCTG 962
 Db 267 PheThrTyHisValLysAsnArgGluGlnAlaTyrglnAsnAspAsnLysThrGlyLeu 286
 QY 963 AATGAAGAAATAACAACACTGACCTGATCTCTGAGAAATATTACCTCTTAAAGAGG 1022
 Db 287 LysLysGluThrLysAsnThrAspLeuIleSerGluLysTyrrIleLeuLysLysGly 306
 QY 1023 GAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTCAAACTGTTCACCAATCAATACGTT 1082
 Db 307 GluLysProTyrrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrrVal 326
 QY 1083 GATGTCGATACCAACGAATTGTCTGAGAGAGCTCTTAAACAGCTAGCGAACGTAAC 1142
 Db 327 AspValAspThrLysAlaLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
 QY 1143 TTAGACTTCAGAGATTATACGATCTCTGATAGGCTTAAGCTTAACACTCTTCAACAATCTC 1202
 Db 347 LeuAspPheArgAspLeuTyrrAspProArgAspLysAlaLysLeuLeuTyrrAsnAsnLeu 366

```

QY 1203 GATGCTTTTGGTATTATGGAATACCTTAACCTGAAAGTAGAGGATAATCAGATGAC 1262
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
367 AspAlaPheAspIleMetAspTyrThrLeuThrGlyLysValGluaspAsnHisAspLys 386
QY 1263 ACCAACCGTATCATACCGTTTATGCGCAAGCGACCCGAGGAGGAGAAATGCTAGCTAT 1322
Db ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
387 AsnAsnArgValValThrValTyrMetGlyLysArgProLysGlyAlaLysGlySerTyr 406
QY 1323 CATTTAGCC 1331
Db ||||| ||||| ||||| |||||
407 HisLeuAla 409

```

Search completed: January 28, 2006, 02:33:51
Job time : 195.003 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2006, 10:53:18 ; Search time 9059.25 Seconds
(without alignments)
11181.386 Million cell updates/sec

Title: US-09-940-235-11
Perfect score: 1782
Sequence: 1 tcgcttcacgttcgtcg.....ataacctaacgacaaataa 1782

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_in.*
- 3: gb_env.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pr.*
- 9: gb_ro.*
- 10: gb_ses.*
- 11: gb_sy.*
- 12: gb_un.*
- 13: gb_vi.*
- 14: gb_htg.*
- 15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1245.8	69.9	2385	6	AR143998 Sequence
2	1244.4	69.8	2568	1	STRSKC
3	1244.4	69.8	2568	6	A04926 S.equisimil
4	1244.4	69.8	8931	1	SEDEXB
5	1240.4	69.6	1242	6	AR144000 Sequence
6	1239.6	69.6	2568	6	E00522 DNA fragmen
7	1237	69.4	1401	6	I05204 Sequence 5
8	1230.4	69.0	2566	6	AR068768
9	1223.4	68.7	1458	6	A20027 SEQ ID NO:
10	1223.4	68.7	1458	6	I13215 Sequence 42
11	1222.6	68.6	1245	1	S46536
12	1222.6	68.6	7057	6	CQ797820
13	1220.4	68.5	1335	6	A20006 SEQ ID NO:
14	1220.4	68.5	1335	6	I13194 Sequence 14
15	1220.4	68.5	1512	6	A20016 SEQ ID NO:
16	1220.4	68.5	1512	6	I13204
17	1220.2	68.5	2589	6	I13209 Sequence 34
18	1219.4	68.4	1257	6	A20015 SEQ ID NO:

19	1219.4	68.4	1257	6	I13203
20	1219.4	68.4	1317	6	A20009
21	1219.4	68.4	1317	6	I13197 Sequence 18
22	1216.4	68.3	1467	6	A20030 SEQ ID NO:
23	1216.4	68.3	1467	6	I13218
24	1216.4	68.3	2588	6	A20021
25	1212.4	68.0	1473	1	SGSKG
26	1211.4	68.0	1407	6	E01413
27	1210.8	67.9	1311	1	AY368335
28	1184	66.4	1209	6	AR175891
29	1184	66.4	1209	6	AX030315
30	1181.4	66.3	1245	6	AR175892
31	1181.4	66.3	1245	6	AX030316
32	1095.6	61.5	50354	1	AE014169
33	1095.6	61.5	110000	1	BA000034_16
34	1095.6	61.5	110000	1	BA000034_17
35	1094	61.4	1323	1	AY234129
36	1094	61.4	1323	1	AY234130
37	1094	61.4	1323	1	AY234137
38	1094	61.4	110000	1	CP000003_16
39	1093.6	61.4	1122	6	AR175893
40	1093.6	61.4	1122	6	AX030317
41	1093.6	61.4	1158	6	AR175894
42	1093.6	61.4	1158	6	AX030318
43	1089.2	61.1	1323	1	AY234136
44	1081.2	60.7	1323	1	AY234128
45	1081.2	60.7	1323	6	CQ872087

ALIGNMENTS

RESULT 1
LOCUS AR143998 2385 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 1 from patent US 6210667.
ACCESSION AR143998
VERSION AR143998.1 GI:15105865
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2385)
AUTHORS Reed G.L.
TITLE Bacterial fibrin-dependent plasminogen activator
JOURNAL Patent: US 6210667-A 1 03-APR-2001;
FEATURES Location/Qualifiers
source
1..2385
/organism="unknown"
/mol_type="unassigned DNA"

Query Match	69.9%;	Score 1245.8;	DB 6;	Length 2385;
Best Local Similarity	99.4%;	Pred. No. 0;		
Matches 1250;	Conservative 0;	Mismatches 7;	Indels 0;	Gaps 0;
Qy	523	TTCCACCATGTTGCTATTGCTGGACCTGAGTGCTGTAGACCTCCATCTGTCAACAAC	582	
Db	1129	TCCATCGAGGTAGGATTGCTGGACCTGAGTGCTGTAGACCTCCATCTGTCAACAAC	1188	
Qy	583	AGCCAATGCTTGTAGCTTGTCTGTTGAGGGGAGCATCAAGACATTAGTCTT	642	
Db	1189	AGCCAATGCTTGTAGCTTGTCTGTTGAGGGGAGCATCAAGACATTAGTCTT	1248	
Qy	643	AAATTTTGAATTCGATCTAAACATCAGACCTCTCATGAGGAAAGACAGACAAAGC	702	
Db	1249	AAATTTTGAATTCGATCTAAACATCAGACCTCTCATGAGGAAAGACAGACAAAGC	1308	
Qy	703	TTAAGTCCAAAATCAAAACCATTTTGTCTACTGATAGTCGCGCATGTGCACATAAACTTGA	762	
Db	1309	TTAAGTCCAAAATCAAAACCATTTTGTCTACTGATAGTCGCGCATGTGCACATAAACTTGA	1368	
Qy	763	AAAGCTGACTTACTTAAAGGCTATTCAAGACAAATTGATCGCTAACGTCCACAGTAACGAC	822	

Db	1369	AAAGCTGACTTACTAAAGGCTATTCAAGAACAAATTGATCGCTAAAGCTCCACAGTAACGAC	1428		
Qy	823	GACTACTTTGAGTCAATTTTGAAGCGATGCAACCATTTACTGATCGAAACGGCAAG	892		
Db	1429	GACTACTTTGAGTCAATTTTGAAGCGATGCAACCATTTACTGATCGAAACGGCAAG	1488		
Qy	883	GTCTACTTTGCTGACAAAGATGGTTCGGTAACCTTGCCGACCCCAACCTGTCGAAGAATTT	942		
Db	1489	GTCTACTTTGCTGACAAAGATGGTTCGGTAACCTTGCCGACCCCAACCTGTCGAAGAATTT	1548		
Qy	943	TTGTAAAGCGGACATGTGCGCGTTAGACCATATAAAGAAAAACCAATACAAAACCAAGCG	1002		
Db	1549	TTGTAAAGCGGACATGTGCGCGTTAGACCATATAAAGAAAAACCAATACAAAACCAAGCG	1608		
Qy	1003	AAATCTGTTGATGTGGAATATATCTGACAGTTTACTTCCCTTAAACCTGATGACGATTTTC	1062		
Db	1609	AAATCTGTTGATGTGGAATATATCTGACAGTTTACTTCCCTTAAACCTGATGACGATTTTC	1668		
Qy	1063	AGACCAGGTCTCAAGACTAAGCTATTGAAACACACTAGCTATCGGTGACACCATCACA	1122		
Db	1669	AGACCAGGTCTCAAGACTAAGCTATTGAAACACACTAGCTATCGGTGACACCATCACA	1728		
Qy	1123	TCTCAAGAAATTAAGTCAAGCACAAGCAATTTTAAACAAAACCAACCCAGGCTATACG	1182		
Db	1729	TCTCAAGAAATTAAGTCAAGCACAAGCAATTTTAAACAAAACCAACCCAGGCTATACG	1788		
Qy	1183	ATTATGAACGATGCTCTCAATCGTCACTCATGACAAATGACATTTTCCGTAGCATTTTA	1242		
Db	1789	ATTATGAACGATGCTCTCAATCGTCACTCATGACAAATGACATTTTCCGTAGCATTTTA	1848		
Qy	1243	CCAAATGGATCAAGATTTTACTTACCGTGTAAAAATCGGGAACACAGCTTATAGGATCAAT	1302		
Db	1849	CCAAATGGATCAAGATTTTACTTACCGTGTAAAAATCGGGAACACAGCTTATAGGATCAAT	1908		
Qy	1303	AAAAAATCTGGTCTGAATGAAGAAATAAACAACACTGATCTCTGAGAAATATTAC	1362		
Db	1909	AAAAAATCTGGTCTGAATGAAGAAATAAACAACACTGATCTCTGAGAAATATTAC	1968		
Qy	1363	GTCTTTAAAAAAGGGGAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACATGTTTC	1422		
Db	1969	GTCTTTAAAAAAGGGGAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACATGTTTC	2028		
Qy	1423	ACCATCAATAGCTTGATGTCATACCAACGAATGCTTAAAGAGTGACAGCTCTTTAACA	1482		
Db	2029	ACCATCAATAGCTTGATGTCATACCAACGAATGCTTAAAGAGTGACAGCTCTTTAACA	2088		
Qy	1483	GCTAGCGAAACGTAACCTTAGACTTTCAGAGATTTTATACGATCCTCGTGAAGGCTAAACTA	1542		
Db	2089	GCTAGCGAAACGTAACCTTAGACTTTCAGAGATTTTATACGATCCTCGTGAAGGCTAAACTA	2148		
Qy	1543	CTCTACAACAAATCTCGATGCTTTTGGTATTATGGACTATACCTTAACTGGAAGTAGAG	1602		
Db	2149	CTCTACAACAAATCTCGATGCTTTTGGTATTATGGACTATACCTTAACTGGAAGTAGAG	2208		
Qy	1603	GATATACAGATGACACCAACCGTATCATACCGTTTATATGGGGAAGCGACCCGAAGGA	1662		
Db	2209	GATATACAGATGACACCAACCGTATCATACCGTTTATATGGGGAAGCGACCCGAAGGA	2268		
Qy	1663	GAGAAATGCTAGCTATCATTTAGCTTATGAAATGATCGTTTATACCGAAGAACCGAGAA	1722		
Db	2269	GAGAAATGCTAGCTATCATTTAGCTTATGAAATGATCGTTTATACCGAAGAACCGAGAA	2328		
Qy	1723	GTTTACAGTACTCTGGTTTATACAGGACACCTTATACCTGATAACCCCTAACGACAAA	1779		
Db	2329	GTTTACAGTACTCTGGTTTATACAGGACACCTTATACCTGATAACCCCTAACGACAAA	2385		
RESULT 2					
STRSKC					
LOCUS					
DEFINITION	Streptococcus equisimilis (H46A) streptokinase gene, complete cds.				
ACCESSION	K02986				

QY	837	CATTGATTTTGCAGCGATGCAACCAATTA	CTGATCGAAACGGCAAGTCTACTTTGCTGA	896
Db	1196	CATTGATTTTGCAGCGATGCAACCAATTA	CTGATCGAAACGGCAAGTCTACTTTGCTGA	1255
QY	897	CAAGATGTTGGTAACTTGGCGCCCAACCTG	TCCAGAAATTTTGGCTTAAGCGACA	956
Db	1256	CAAGATGTTGGTAACTTGGCGCCCAACCTG	TCCAGAAATTTTGGCTTAAGCGACA	1315
QY	957	TGTGCGCGTTAGACCATATAAAGAAAAA	CCAAATACAAACCAACCGGAAATCTGTTGATGT	1016
Db	1316	TGTGCGCGTTAGACCATATAAAGAAAAA	CCAAATACAAACCAACCGGAAATCTGTTGATGT	1375
QY	1017	GGAATATACGTACAGTTTACTCCCTTAAAC	CCCTGATGACGATTTTCAGACCGCTCTCAA	1076
Db	1376	GGAATATACGTACAGTTTACTCCCTTAAAC	CCCTGATGACGATTTTCAGACCGCTCTCAA	1435
QY	1077	AGTACTAAGCTATTGAAACACATAGCTAT	CGGTGACACCATCTCAAGAAATTA	1136
Db	1436	AGTACTAAGCTATTGAAACACATAGCTAT	CGGTGACACCATCTCAAGAAATTA	1495
QY	1137	AGCTCAAGCACAAAGCAATTTTAAACAA	AAAAACCAACCGGCTTATACGATTTATGAACGTGA	1196
Db	1496	AGCTCAAGCACAAAGCAATTTTAAACAA	AAAAACCAACCGGCTTATACGATTTATGAACGTGA	1555
QY	1197	CTCCTCAATCGTCACTCATGACAATGACA	TATTTCCGTTACGATTTTACCAATGGATCAAGA	1256
Db	1556	CTCCTCAATCGTCACTCATGACAATGACA	TATTTCCGTTACGATTTTACCAATGGATCAAGA	1615
QY	1257	GTTTACTTACCGTGTAAATCGGAAACA	GCTTATAGGATCAATAAAAATCTGGTCT	1316
Db	1616	GTTTACTTACCGTGTAAATCGGAAACA	GCTTATAGGATCAATAAAAATCTGGTCT	1675
QY	1317	GAATGAAGAAATAAACAACACTGACCTG	ATCTCTGAGAAATATTAACGCTCTTAAAAAAGG	1376
Db	1676	GAATGAAGAAATAAACAACACTGACCTG	ATCTCTGAGAAATATTAACGCTCTTAAAAAAGG	1735
QY	1377	GGAAAGCGGTATGATTCCTTTGATCGC	AGTCACTTGAACCTGTTACCAATCAATACGT	1436
Db	1736	GGAAAGCGGTATGATTCCTTTGATCGC	AGTCACTTGAACCTGTTACCAATCAATACGT	1795
QY	1437	TGATGTCGATACCAACGAATTTGCTAA	AAAAAGTAGAGAGCTCTTAACAGCTAGCGAACGTAA	1496
Db	1796	TGATGTCGATACCAACGAATTTGCTAA	AAAAAGTAGAGAGCTCTTAACAGCTAGCGAACGTAA	1855
QY	1497	CTTAGACTTCAGAGATTTATACGATCCT	CGTGTATAGGCTTAAACTACTCTACACAAATCT	1556
Db	1856	CTTAGACTTCAGAGATTTATACGATCCT	CGTGTATAGGCTTAAACTACTCTACACAAATCT	1915
QY	1557	CGATGCTTTTGGTATTATGGACTATAC	CTTTAACTGGAAAAAGTAGAGATTAATCACGATGA	1616
Db	1916	CGATGCTTTTGGTATTATGGACTATAC	CTTTAACTGGAAAAAGTAGAGATTAATCACGATGA	1975
QY	1617	CACCAACCGTATCATAAACCGTTTATAT	GGGCAAGCGACCCGAGGAGAGATGCTAGCTA	1676
Db	1976	CACCAACCGTATCATAAACCGTTTATAT	GGGCAAGCGACCCGAGGAGAGATGCTAGCTA	2035
QY	1677	TCATTTAGCTTATGATAAGATCGTTAT	ATACCGAAGAAACGAGAAATTTACAGCTACCT	1736
Db	2036	TCATTTAGCTTATGATAAGATCGTTAT	ATACCGAAGAAACGAGAAATTTACAGCTACCT	2095
QY	1737	GGGTTATACAGGACACCTTATACCTGAT	AACCCCTTAACGACAAATAA	1782
Db	2096	GGGTTATACAGGACACCTTATACCTGAT	AACCCCTTAACGACAAATAA	2141

RESULT 3
A04926
LOCUS A04926 2568 bp DNA linear PAT 15-JUL-1993
DEFINITION S.equisimilis skc gene for streptokinase.
ACCESSION A04926
VERSION A04926.1 GI:412219
KEYWORDS streptokinase.
SOURCE Streptococcus dyegalactiae subsp. equisimilis

ORGANISM	Streptococcus dysgalactiae subsp. equisimilis									
	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;									
REFERENCE	Streptococcus.									
AUTHORS	1 (bases 1 to 2568)									
TITLE	Ferretti, J.J. and Malke, H.									
JOURNAL	Streptokinase-coding recombinant vectors									
	Patent: EP 0151337-A 1 14-AUG-1985;									
	THE BOARD OF REGENTS OF THE UNIVERSITY OF OKLAHOMA									
FEATURES	Location/Qualifiers									
source	1..2568									
	/organism="Streptococcus dysgalactiae subsp. equisimilis"									
	/mol_type="unassigned DNA"									
	/sub_species="equisimilis"									
	/db_xref="taxon:119602"									
ORIGIN										
Query Match	69.8%; Score 1244.4; DB 6; Length 2568;									
Best Local Similarity	99.9%; Pred. No. 0;									
Matches 1245; Conservative	0; Mismatches 1; Indels 0; Gaps 0;									
QY	537	TATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAA	CAGCAATTTGGTTGT	596						
Db	896	TATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAA	CAGCAATTTGGTTGT	955						
QY	597	TAGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTTCGAAAT	656							
Db	956	TAGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTTCGAAAT	1015							
QY	657	CGATCTAACATCAACGACCTGCTCATGAGGAAAGACAGACGAGCGCTTAAGTCCAAATC	716							
Db	1016	CGATCTAACATCAACGACCTGCTCATGAGGAAAGACAGACGAGCGCTTAAGTCCAAATC	1075							
QY	717	AAAAACCATTTGCTACTGATGCGGCGATGTACATAAACTTGAGAAAGCTGACTTACT	776							
Db	1076	AAAAACCATTTGCTACTGATGCGGCGATGTACATAAACTTGAGAAAGCTGACTTACT	1135							
QY	777	AAAGGCTATTCAAGAAACAATTTGATCGCTAACGTCACAGTAACGACGACTACTTTGAGGT	836							
Db	1136	AAAGGCTATTCAAGAAACAATTTGATCGCTAACGTCACAGTAACGACGACTACTTTGAGGT	1195							
QY	837	CATTGATTTTGCAGCGATGCAACCAATTA	CTGATCGAAACGGCAAGGCTCTACTTTGCTGA	896						
Db	1196	CATTGATTTTGCAGCGATGCAACCAATTA	CTGATCGAAACGGCAAGGCTCTACTTTGCTGA	1255						
QY	897	CAAAAGATGTTGCGTTAACTTGGCGCCCAACCTGTCCAAAGAAATTTTTCGCTAAGCGGACA	956							
Db	1256	CAAAAGATGTTGCGTTAACTTGGCGCCCAACCTGTCCAAAGAAATTTTTCGCTAAGCGGACA	1315							
QY	957	TGTGCGCGTTAGACCATATAAAGAAAAA	CCAAATACAAAACCAAGCGAAATCTGTTGATGT	1016						
Db	1316	TGTGCGCGTTAGACCATATAAAGAAAAA	CCAAATACAAAACCAAGCGAAATCTGTTGATGT	1375						
QY	1017	GGAATATACGTACAGTTTACTCCCTTAAACCTGTGATGCGAATTTTCAGAACGAGGCTCAA	1076							
Db	1376	GGAATATACGTACAGTTTACTCCCTTAAACCTGTGATGCGAATTTTCAGAACGAGGCTCAA	1435							
QY	1077	AGATCTAAGCTATTGAAAAACATAGCTATCGGTGACACCATCTCAAGAAATTA	1136							
Db	1436	AGATCTAAGCTATTGAAAAACATAGCTATCGGTGACACCATCTCAAGAAATTA	1495							
QY	1137	AGCTCAAGCACAAAGCAATTTTAAACAAAACCAACCGAGGCTATACGATTTATGAACGTGA	1196							
Db	1496	AGCTCAAGCACAAAGCAATTTTAAACAAAACCAACCGAGGCTATACGATTTATGAACGTGA	1555							
QY	1197	CTCCTCAATCGTCACTCATGACAATGACAATTTTCGGTACGATTTTACCAATGGATCAAGA	1256							
Db	1556	CTCCTCAATCGTCACTCATGACAATGACAATTTTCGGTACGATTTTACCAATGGATCAAGA	1615							
QY	1257	GTTTTACTTACCGTGTAAAAATCGGGAAACAAGCTTATAGGATCAATAAAAATCTGGTCT	1316							
Db	1616	GTTTTACTTACCGTGTAAAAATCGGGAAACAAGCTTATAGGATCAATAAAAATCTGGTCT	1675							
QY	1317	GAATGAAGAAAAATAACAACACTGACCTGATCTCTGAGAAATATTAACGCTCTTAAAAAAGG	1376							


```

CDS
complement(3262..4110)
/gene="lrp"
/codon_start=1
/transl_table=11
/product="leucine rich protein"
/protein_id="CAA51350.1"
/db_xref="GI:407878"
/db_xref="GOA:Q54087"
/db_xref="InterPro:IPR002197"
/db_xref="InterPro:IPR008931"
/db_xref="UniProt/Swiss-Prot:Q54087"
/translation="MELKDYFPEMVQPHPLGDKWVSVEGQYVHFPPKSCISEKER
LLEVLGQYELVQPGSPQWRYLLDHQGNPPQLFQTSQFIYLNHQQVLPADLVLLQ
QMAGLEVLPISSTQATFLCQATSIKVLRSLEGLPTLESDFGLATFMFVGNAMTQ
VAAGTREFEECCQLLTAYLKQKSGKLTFAEVLMSILSHQSPALTRQPHQFLN
PQSDMADVVAHWLSEHNLVQTAQRLYIHRNSLYQKLDKFAQQSGSLHLKQLDLDLAPAY
LFLLYK"
complement(3337..3423)
/gene="lrp"
/feature="leucine zipper motif"
complement(4117..4120)
/gene="lrp"
complement(4141..4146)
/gene="lrp"
complement(4148..4149)
/gene="lrp"
/feature="TG motif"
4172..4368
/feature="static DNA bending locus"
4392..5837
/gene="skc"
4392..4393
/gene="skc"
/feature="TG motif"
4395..4400
/gene="skc"
4401..4406
/feature="skc"
4428..4431
/feature="alternative"
4439..5761
/gene="skc"
/feature="skc"
/codon_start=1
/transl_table=11
/product="streptokinase"
/protein_id="CAA51351.1"
/db_xref="GI:407879"
/db_xref="GOA:Q54088"
/db_xref="InterPro:IPR004093"
/db_xref="InterPro:IPR008124"
/db_xref="UniProt/Swiss-Prot:P00779"
/translation="MKNYLSFGMFAALLFALTFTGVNSVQAIAGPEWLLDRPVSNNSQL
VVSVAQVETGNTQDLSLKFEDLTSRPAHGKTEQGLSPKSPFATDSGAMSHKLEK
ADLLKAOEQLIYVNSNDYDEVIDFASDATITDRNGKVPYADKXGVTLPQVQE
FLLSGHVVRPKERQIONAKSVDEVTQPTLPNPDPRGLKDTLLKTLAIGD
TITSELLAQOQSLKHPGTIYERDSSIYVHDNDIFRTILPMQEFYRVKREQ
AYRNKSGLENEINTDLISIKYVVKKGEKPPDFDRSHLKLFTIKYVDVDTNELL
KSEQLLTASRENDLFRDLYDPRDKAKLLYNNLDAFGIMDYTLTKGVEDNHDNTRIIT
VYMGKRPGEANSHYLAVDKORYTBEEREVSYLYRTGTPIPDNPNDK"
4439..4516
/gene="skc"
5796..5837
/gene="skc"
complement(5859..6302)
/feature="unnamed protein product; ORF1"
/codon_start=1
/transl_table=11
/protein_id="CAA51352.1"
/db_xref="GI:407880"
/db_xref="GOA:Q54088"
/db_xref="InterPro:IPR003732"
/db_xref="UniProt/Swiss-Prot:Q54088"

misc_feature
/gene="lrp"
/feature="leucine zipper motif"
complement(4117..4120)
/gene="lrp"
complement(4141..4146)
/gene="lrp"
complement(4148..4149)
/gene="lrp"
/feature="TG motif"
4172..4368
/feature="static DNA bending locus"
4392..5837
/gene="skc"
4392..4393
/gene="skc"
/feature="TG motif"
4395..4400
/gene="skc"
4401..4406
/feature="skc"
4428..4431
/feature="alternative"
4439..5761
/gene="skc"
/feature="skc"
/codon_start=1
/transl_table=11
/product="streptokinase"
/protein_id="CAA51351.1"
/db_xref="GI:407879"
/db_xref="GOA:Q54088"
/db_xref="InterPro:IPR004093"
/db_xref="InterPro:IPR008124"
/db_xref="UniProt/Swiss-Prot:P00779"
/translation="MKNYLSFGMFAALLFALTFTGVNSVQAIAGPEWLLDRPVSNNSQL
VVSVAQVETGNTQDLSLKFEDLTSRPAHGKTEQGLSPKSPFATDSGAMSHKLEK
ADLLKAOEQLIYVNSNDYDEVIDFASDATITDRNGKVPYADKXGVTLPQVQE
FLLSGHVVRPKERQIONAKSVDEVTQPTLPNPDPRGLKDTLLKTLAIGD
TITSELLAQOQSLKHPGTIYERDSSIYVHDNDIFRTILPMQEFYRVKREQ
AYRNKSGLENEINTDLISIKYVVKKGEKPPDFDRSHLKLFTIKYVDVDTNELL
KSEQLLTASRENDLFRDLYDPRDKAKLLYNNLDAFGIMDYTLTKGVEDNHDNTRIIT
VYMGKRPGEANSHYLAVDKORYTBEEREVSYLYRTGTPIPDNPNDK"
4439..4516
/gene="skc"
5796..5837
/gene="skc"
complement(5859..6302)
/feature="unnamed protein product; ORF1"
/codon_start=1
/transl_table=11
/protein_id="CAA51352.1"
/db_xref="GI:407880"
/db_xref="GOA:Q54088"
/db_xref="InterPro:IPR003732"
/db_xref="UniProt/Swiss-Prot:Q54088"

sig_peptide
/gene="skc"
5796..5837
/gene="skc"
terminator
CDS
complement(5859..6302)
/feature="unnamed protein product; ORF1"
/codon_start=1
/transl_table=11
/protein_id="CAA51352.1"
/db_xref="GI:407880"
/db_xref="GOA:Q54088"
/db_xref="InterPro:IPR003732"
/db_xref="UniProt/Swiss-Prot:Q54088"

Query Match 69.8%; Score 1244.4; DB 1; Length 8931;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1245; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 537 TATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTCTCAACACAGCCCAATTGGTGT 596
Db 4516 TATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTCTCAACACAGCCCAATTGGTGT 4575
Qy 597 TAGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAAT 656
Db 4576 TAGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAAT 4635
Qy 657 CGATCTAACATCAGCAGCTGCTCATGAGGAAAGACAGAGCAAGCTTTAAGTCCAAATC 716
Db 4636 CGATCTAACATCAGCAGCTGCTCATGAGGAAAGACAGAGCAAGCTTTAAGTCCAAATC 4695
Qy 717 AAACCAATTTGCTACTGATAGTGGCGGATGTCATATAAATCTTGAAGAGTGACTTACT 776
Db 4696 AAACCAATTTGCTACTGATAGTGGCGGATGTCATATAAATCTTGAAGAGTGACTTACT 4755
Qy 777 AAAGGCTATTCAAGAACAAATTTGATCGCTTAAGCTCCACAGTAAACACGACTACTTTGAGGT 836
Db 4756 AAAGGCTATTCAAGAACAAATTTGATCGCTTAAGCTCCACAGTAAACACGACTACTTTGAGGT 4815
Qy 837 CATTTGATTTTGAAGCGATGCAACCATTTACTGTGCGAACCGCAAGGTCTACTTTGCTGA 896
Db 4816 CATTTGATTTTGAAGCGATGCAACCATTTACTGTGCGAACCGCAAGGTCTACTTTGCTGA 4875
Qy 897 CAAAGATGGTTCCGTAACTTCCGACCGCCAACTGTCCAGAAATTTTGTCAAGCGGACA 956
Db 4876 CAAAGATGGTTCCGTAACTTCCGACCGCCAACTGTCCAGAAATTTTGTCAAGCGGACA 4935
Qy 957 TGTGCGGTTTAGACCATATATAAGAAAAACCAATACAAACCAAGCAAACTCTGTTGATGT 1016
Db 4936 TGTGCGGTTTAGACCATATATAAGAAAAACCAATACAAACCAAGCAAACTCTGTTGATGT 4995
Qy 1017 GGAATATCTGTACAGTTTACTCCCTTAAACCCCTGTGATGACGATTTTCAAGCAGGTCTCAA 1076
Db 4996 GGAATATCTGTACAGTTTACTCCCTTAAACCCCTGTGATGACGATTTTCAAGCAGGTCTCAA 5055
Qy 1077 AGATATCAAGTATTGAAAAACATAGTCTATCGGTGACACCATCATCATCTCAAGAATTACT 1136
Db 5056 AGATATCAAGTATTGAAAAACATAGTCTATCGGTGACACCATCATCATCTCAAGAATTACT 5115
Qy 1137 AGCTCAAGCAAAAGCAATTTTAAACAAACCCAGGCTATACGATTTATGAACTGA 1196
Db 5116 AGCTCAAGCAAAAGCAATTTTAAACAAACCCAGGCTATACGATTTATGAACTGA 5175
Qy 1197 CTCTCAATCGTCACTCATGCAATGACATTTTCCGTACGATTTTACCAATGATCAAGA 1256
Db 5176 CTCTCAATCGTCACTCATGCAATGACATTTTCCGTACGATTTTACCAATGATCAAGA 5235
Qy 1257 GTTTACTTACCGTTTAAAAAATCGGAAACAGCTTATAGGATCAATAAAAAATCTGCTCT 1316
Db 5236 GTTTACTTACCGTTTAAAAAATCGGAAACAGCTTATAGGATCAATAAAAAATCTGCTCT 5295
Qy 1317 GAATGAGAAATAAACAACACTGACCTGATCTCTGAGAAATATATACGTCCTTAAAAAAGG 1376

```


Db 5296 GAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGG 5355
 Qy 1377 GGAAAGCCGATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCCACCATCAATACGT 1436
 Db 5356 GGAAAGCCGATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCCACCATCAATACGT 5415
 Qy 1437 TGATGTCGATACCAACGAATTCGTAAGAGTGAAGCTCTTAAACAGCTAGCGAAGCTAA 1496
 Db 5416 TGATGTCGATACCAACGAATTCGTAAGAGTGAAGCTCTTAAACAGCTAGCGAAGCTAA 5475
 Qy 1497 CTTAGACTTCAGAGATTTATACGATCCTCGTGAAGAGTGAAGCTCTTAAACAGCTAGCGAAGCTAA 1556
 Db 5476 CTTAGACTTCAGAGATTTATACGATCCTCGTGAAGAGTGAAGCTCTTAAACAGCTAGCGAAGCTAA 5535
 Qy 1557 CGATGCTTTTGGTATATGATGCTATACCTTAACTGGAAGAGTGAAGCTCTTAAACAGCTAGCGAAGCTAA 1616
 Db 5536 CGATGCTTTTGGTATATGATGCTATACCTTAACTGGAAGAGTGAAGCTCTTAAACAGCTAGCGAAGCTAA 5595
 Qy 1617 CACCAACCGTATCATAAACCGTTTATATGGCAAGCGACCCGAGGAGAGATGCTAGCTA 1676
 Db 5596 CACCAACCGTATCATAAACCGTTTATATGGCAAGCGACCCGAGGAGAGATGCTAGCTA 5655
 Qy 1677 TCATTTAGCCTATGATAAAGATCGTTTATACCGAAGAGAACCGTAAGCTTTACAGCTACCT 1736
 Db 5656 TCATTTAGCCTATGATAAAGATCGTTTATACCGAAGAGAACCGTAAGCTTTACAGCTACCT 5715
 Qy 1737 GCGTTATACAGGACACCTATACCTGATTAACCTTAACCGTAAGAACCGTAAGCTTTACAGCTACCT 1782
 Db 5716 GCGTTATACAGGACACCTATACCTGATTAACCTTAACCGTAAGAACCGTAAGCTTTACAGCTACCT 5761

RESULT 5
 ARI144000
 LOCUS ARI144000 1242 bp DNA linear PAT 08-AUG-2001
 DEFINITION Sequence 5 from patent US 6210667.
 ACCESSION ARI144000
 VERSION ARI144000.1 GI:15105867
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1242)
 Unclassified.
 AUTHORS Reed, G. L.
 TITLE Bacterial fibrin-dependent plasminogen activator
 JOURNAL Patent: US 6210667-A 5 03-APR-2001;
 FEATURES
 source 1..1242
 /organism="unknown"
 /mol_type="unassigned DNA"

ORIGIN
 Query Match 69.6%; Score 1240.4; DB 6; Length 1242;
 Best Local Similarity 99.9%; Pred. No. 2.8e-313;
 Matches 1241; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 538 ATTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAGCCCAATTTGGTTGTT 597
 Db 1 ATTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAGCCCAATTTAGTTGTT 60
 Qy 598 AGCGTTGCTGTTGAGGGGACGATCAAGCATTTAGTCTTAAATTTTGGAAATC 657
 Db 61 AGCGTTGCTGTTGAGGGGACGATCAAGCATTTAGTCTTAAATTTTGGAAATC 120
 Qy 658 GATCTTAACATCACGACCTGCTCATGGAGGAAAGACAGAGCGCTTAAAGTCCAAATCA 717
 Db 121 GATCTTAACATCACGACCTGCTCATGGAGGAAAGACAGAGCGCTTAAAGTCCAAATCA 180
 Qy 718 AAACCATTTGCTGATGATGTCGCGATGTCAATAAATTGAGAAAGCTGACTTACTA 777
 Db 181 AAACCATTTGCTGATGATGTCGCGATGTCAATAAATTGAGAAAGCTGACTTACTA 240
 Qy 778 AAGGCTATTCAAGCAATTTGATCGTACGTTCCACGTAACGACACTTCTTGGTCTC 837

Db 241 AAGGCTATTCAAGAAACAATTTGATCGCTTAAACGTCACAGTAACGACGACTACTTTGAGTCT 300
 Qy 838 ATTGATTTTGCAGCGATGCAACCATTTACTGATCGAAGCGCAAGGCTTACTTTGCTGAC 897
 Db 301 ATTGATTTTGCAGCGATGCAACCATTTACTGATCGAAGCGCAAGGCTTACTTTGCTGAC 360
 Qy 898 AAGATGTTTCCGTTAACTTTGCCGACCAACCTCTCCAAAGATTTTGTCTAAGCGGACAT 957
 Db 361 AAGATGTTTCCGTTAACTTTGCCGACCAACCTCTCCAAAGATTTTGTCTAAGCGGACAT 420
 Qy 958 GTGCGGTTTGAACCATTAAGAAAAACCAATACAAAAACCAAGCGAAATCTGTTGATGTG 1017
 Db 421 GTGCGGTTTGAACCATTAAGAAAAACCAATACAAAAACCAAGCGAAATCTGTTGATGTG 480
 Qy 1018 GAATATATCTGACAGTTTACTTCCCTTAAACCGTGTGATGACGATTTTTCAGACGAGTCTCAA 1077
 Db 481 GAATATATCTGACAGTTTACTTCCCTTAAACCGTGTGATGACGATTTTTCAGACGAGTCTCAA 540
 Qy 1078 GATATTAAGCTTTTGAACCACTAGCTATCGGTGACACCATCATCTCAAGAAATTACTA 1137
 Db 541 GATATTAAGCTTTTGAACCACTAGCTATCGGTGACACCATCATCTCAAGAAATTACTA 600
 Qy 1138 GCTCAAGCACAAGCAATTTTAAACAAAAACCAAGCGCTATACGATTTTATSAACGTGAC 1197
 Db 601 GCTCAAGCACAAGCAATTTTAAACAAAAACCAAGCGCTATACGATTTTATSAACGTGAC 660
 Qy 1198 TCCTCAATGCTCACTCATGACCAATTTTCCGTACGATTTTACCAATGATCAAGAG 1257
 Db 661 TCCTCAATGCTCACTCATGACCAATTTTCCGTACGATTTTACCAATGATCAAGAG 720
 Qy 1258 TTTTACTTACCGTGTAAATAATCGGAAACAGCTTTATAGGATCAATAAATAATCTGGTCTG 1317
 Db 721 TTTTACTTACCGTGTAAATAATCGGAAACAGCTTTATAGGATCAATAAATAATCTGGTCTG 780
 Qy 1318 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGCTCTTAAAAAGGG 1377
 Db 781 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGCTCTTAAAAAGGG 840
 Qy 1378 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTTGAAACCTGTTCAACATCAATACGTT 1437
 Db 841 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTTGAAACCTGTTCAACATCAATACGTT 900
 Qy 1438 GATGTCGATACCAACGAATTTGCTTAAAGTGAGGAGCTCTTAAACAGCTAGGAGCGTAAC 1497
 Db 901 GATGTCGATACCAACGAATTTGCTTAAAGTGAGGAGCTCTTAAACAGCTAGGAGCGTAAC 960
 Qy 1498 TTAGACTTCAGAGATTTATACGATCCTCGTGAAGGCTAAACTTACTCTACCAATCTC 1557
 Db 961 TTAGACTTCAGAGATTTATACGATCCTCGTGAAGGCTAAACTTACTCTACCAATCTC 1020
 Qy 1558 GATGCTTTTGGTATTTAGGACTATACCTTAACTGGAAGAGTAGAGGATAATCACGATGAC 1617
 Db 1021 GATGCTTTTGGTATTTAGGACTATACCTTAACTGGAAGAGTAGAGGATAATCACGATGAC 1080
 Qy 1618 ACCAAGCGTATACCGTTTATATGGGCAAGCGACCCGAGGAGAGATGCTAGCTAT 1677
 Db 1081 ACCAAGCGTATACCGTTTATATGGGCAAGCGACCCGAGGAGAGATGCTAGCTAT 1140
 Qy 1678 CATTTAGCCTTATGATAAAGATCGTTTATACCGAAGAGAACGAGAAATTTACAGCTACCTG 1737
 Db 1141 CATTTAGCCTTATGATAAAGATCGTTTATACCGAAGAGAACGAGAAATTTACAGCTACCTG 1200
 Qy 1738 CGTTATACAGGACACCTATACCTGATAACCCCTAAACGACAAA 1779
 Db 1201 CGTTATACAGGACACCTATACCTGATAACCCCTAAACGACAAA 1242

RESULT 6
 E00522
 LOCUS E00522 2568 bp DNA linear PAT 29-SEP-1997
 DEFINITION DNA fragment of plasmid PMF1 into which DNA encoding streptokinase is inserted.

ACCESSION E00522
 VERSION E00522.1 GI:2168801
 KEYWORDS JP 1985237995-A/1.
 SOURCE Streptococcus dysgalactiae subsp. equisimilis
 ORGANISM Streptococcus dysgalactiae subsp. equisimilis
 Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 Streptococcus.
 1 (bases 1 to 2568)
 REFERENCE Jiyosefu, J.F. and Horusuto, M.
 AUTHORS RECOMBINATION VECTOR FOR PRODUCING STREPTOKINASE
 TITLE Patent: JP 1985237995-A 1 26-NOV-1985;
 JOURNAL PHILLIPS PETROLEUM CO
 COMMENT OS Streptococcus equisimilis
 PN JP 1985237995-A/1
 PD 26-NOV-1985
 PF 09-OCT-1984 JP 1984212403
 PR 10-OCT-1983 DD 83 255523, 02-MAR-1984 US 84 585417 PI
 JIYOSUFU JIEI FURETSUTSUI, HORUSUTO MARUKE
 PC C12N15/00, C12N1/20, C12N9/70, (C12N15/00, C12R1:46), (C12N1/20, PC
 C12R1:19),
 PC (C12N9/70, C12R1:19);
 CC strandedness: Single;
 CC topology: Linear;
 CC hypothetical: No;
 CC anti-sense: No;
 CC *source: strains=Streptococcus equisimilis H46A; CC *source:
 library=streptococcus equisimilis H46 library; CC *source:
 clone=lambda L47 skc clone;
 FH Key Location/Qualifiers
 FT 5'UTR 1..818
 FT sig_peptide 819..896
 FT /product='streptokinase signal peptide' FT
 mat_peptide 897..2138
 FT /product='streptokinase'
 FT CDS 819..2143
 FT /product='streptokinase'
 FT 3'UTR 2142..2569.
 FEATURES Location/Qualifiers
 source 1..2568
 /organism='Streptococcus dysgalactiae subsp. equisimilis'
 /mol_type='genomic DNA'
 /db_xref='taxon:119602'
 ORIGIN
 Query Match 69.6%; Score 1239.6; DB 6; Length 2568;
 Best Local Similarity 99.7%; Pred. No. 5.6e-313;
 Matches 1242; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 537 TATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCACAAACAGCCCAATTGGTTGT 596
 DB 896 TATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCACAAACAGCCCAATTGGTTGT 955
 QY 597 TAGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTCAAAT 656
 DB 956 TAGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTCAAAT 1015
 QY 657 CGATCTAATCATCAGACCTGCTCATGAGGAAAGACAGACGAGGCTTAAGTCCAAATC 716
 DB 1016 CGATCTAATCATCAGACCTGCTCATGAGGAAAGACAGACGAGGCTTAAGTCCAAATC 1075
 QY 717 AAAACCAATTTGCTACTGATAGTGGCGGATGTCACATAAATTGAGAAAGCTGACTTACT 776
 DB 1076 AAAACCAATTTGCTACTGATAGTGGCGGATGTCACATAAATTGAGAAAGCTGACTTACT 1135
 QY 777 AAAGCTATTCAAGAACCAATTGATCGCTAACTCCACAGTACGACGACTCTTTGAGGT 836
 DB 1136 AAAGCTATTCAAGAACCAATTGATCGCTAACTCCACAGTACGACGACTCTTTGAGGT 1195
 QY 837 CATTTGATTTGCAAGCGATGCAACCAATTACTGATCGAAACGGCAAGGTCTACTTTGCTGA 896
 DB 1196 CATTTGATTTGCAAGCGATGCAACCAATTACTGATCGAAACGGCAAGGTCTACTTTGCTGA 1255

QY 897 CAAAGATGGTTGGTAAACCTTGGCCGACCCAACTGTCTCAAGAAATTTTGTCTAAGCGACA 956
 DB 1256 CAAAGATGGTTGGTAAACCTTGGCCGACCCAACTGTCTCAAGAAATTTTGTCTAAGCGACA 1315
 QY 957 TGTGCGGTTTAGACCATATAAGAAAACCAATACAAACCAAGCGAAATCTGTTGATGT 1016
 DB 1316 TGTGCGGTTTAGACCATATAAGAAAACCAATACAAACCAAGCGAAATCTGTTGATGT 1375
 QY 1017 GGAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCAGTCTCAA 1076
 DB 1376 GGAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCAGTCTCAA 1435
 QY 1077 AGATATAAGCTATTGAAAAACATAGCTATCGGTGACACCATCATCTCTCAAGAAATTACT 1136
 DB 1436 AGATATAAGCTATTGAAAAACATAGCTATCGGTGACACCATCATCTCTCAAGAAATTACT 1495
 QY 1137 AGCTCAAGCAAAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTTATGAACGTGA 1196
 DB 1496 AGCTCAAGCAAAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTTATGAACGTGA 1555
 QY 1197 CTCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAAGA 1256
 DB 1556 CTCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAAGA 1615
 QY 1257 GTTTACTTACCGTGTAAATAATCGGAAACCAAGCTTATAGGATCAATAAAAAATCTGCTCT 1316
 DB 1616 GTTTACTTACCGTGTAAATAATCGGAAACCAAGCTTATAGGATCAATAAAAAATCTGCTCT 1675
 QY 1317 GAATGAAGAAATAAACCAACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGG 1376
 DB 1676 GAATGAAGAAATAAACCAACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGG 1735
 QY 1377 GGAAGAGCCGTATGATCCCTTGTATCGAGTCACTTGAACCTGTTACCAATCAATACGT 1436
 DB 1736 GGAAGAGCCGTATGATCCCTTGTATCGAGTCACTTGAACCTGTTACCAATCAATACGT 1795
 QY 1437 TGATGTCGATACCAACGAATTTGCTTAAAGAGTGACAGCTCTTAAACAGCTAGCGAACCTAA 1496
 DB 1796 TGATGTCGATACCAACGAATTTGCTTAAAGAGTGACAGCTCTTAAACAGCTAGCGAACCTAA 1855
 QY 1497 CTTAGACTTCCAGAGATTTTATACGATCTCTGATAGGCTTAACTCTTACCAATCTCT 1556
 DB 1856 CTTAGACTTCCAGAGATTTTATACGATCTCTGATAGGCTTAACTCTTACCAATCTCT 1915
 QY 1557 CGATGCTTTGGTATTATGGAATATACCTTAACTGGAAGAGTAGAGATTAATCAACGATGA 1616
 DB 1916 CGATGCTTTGGTATTATGGAATATACCTTAACTGGAAGAGTAGAGATTAATCAACGATGA 1975
 QY 1617 CACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAGGAGAGATGCTAGCTA 1676
 DB 1976 CACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAGGAGAGATGCTAGCTA 2035
 QY 1677 TCATTTAGCTTATGATAAAGATCGTTATACCGAAGAAACGAGAACTTTACAGCTACCT 1736
 DB 2036 TCATTTAGCTTATGATAAAGATCGTTATACCGAAGAAACGAGAACTTTACAGCTACCT 2095
 QY 1737 GCCTTATACAGGACACCTTATACCTGATAACCTTAACGACAAATAA 1782
 DB 2096 GCCTTATACAGGACACCTTATACCTGATAACCTTAACGACAAATAA 2141
 RESULT 7
 LOCUS 105204
 DEFINITION Sequence 5 from Patent EP 0248227.
 ACCESSION 105204
 VERSION 105204.1 GI:591209
 KEYWORDS Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1401)
 AUTHORS Hagenson, M.J. and Stroman, D.W.

TITLE Yeast production of streptokinase
JOURNAL Patent: EP 0248227-A1 5 09-DEC-1987;
FEATURES Location/Qualifiers
source 1. .1401
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 69.4%; Score 1237; DB 6; Length 1401;
Best Local Similarity 99.6%; Pred. No. 3.1e-312; Indels 0; Gaps 0;
Matches 1240; Conservative 0; Mismatches 5;
Qy 538 ATTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCCAAATTTGGTTGT 597
Db 10 ATTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCCAAATTTAGTTGT 69
Qy 598 AGCGTTGCTGTACTGTTGAGGGGACGAATCAAGACATTTAGTCTTAAATTTTGGAAATC 657
Db 70 AGCGTTGCTGTACTGTTGAGGGGACGAATCAAGACATTTAGTCTTAAATTTTGGAAATC 129
Qy 658 GATCTAAACATCAGACCTGCTCATGAGGAAAGACAGACGAGGCTTAAAGTCCAAATCA 717
Db 130 GATCTAAACATCAGACCTGCTCATGAGGAAAGACAGACGAGGCTTAAAGTCCAAATCA 189
Qy 718 AAACCAATTTGCTAGTAGTGGCGGATGTCTACATAAACTTTGAGAAAGCTGACTTACTA 777
Db 190 AAACCAATTTGCTAGTAGTGGCGGATGTCTACATAAACTTTGAGAAAGCTGACTTACTA 249
Qy 778 AAGGCTATTCAAGAACAAATGATCGCTAACCGTCCACAGTAACGACGACTACTTTGAGGTC 837
Db 250 AAGGCTATTCAAGAACAAATGATCGCTAACCGTCCACAGTAACGACGACTACTTTGAGGTC 309
Qy 838 ATTGATTTTGAAGCGATCAACCATTTACTGTGTAAGCGGCAAGTCTACTTTGCTGAC 897
Db 310 ATTGATTTTGAAGCGATCAACCATTTACTGTGTAAGCGGCAAGTCTACTTTGCTGAC 369
Qy 898 AAAGATGTTTGGTAACTTTCGCGACCCCACTGTCCAAAGAAATTTTGTCTAAGCGGACAT 957
Db 370 AAAGATGTTTGGTAACTTTCGCGACCCCACTGTCCAAAGAAATTTTGTCTAAGCGGACAT 429
Qy 958 GTGGCGGTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTGTATGTG 1017
Db 430 GTGGCGGTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTGTATGTG 489
Qy 1018 GAATATAGTACAGTTTACTCCCTTAAACCTGTATGACGATTTACAGCAGTCTCAAA 1077
Db 490 GAATATAGTACAGTTTACTCCCTTAAACCTGTATGACGATTTTACAGCAGTCTCAAA 549
Qy 1078 GATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATTTACTA 1137
Db 550 GATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATTTACTA 609
Qy 1138 GCTCAAGCACAAGCATTTTAAACAAAAACCCAGGCTATACGATTTTATGAACGTGAC 1197
Db 610 GCTCAAGCACAAGCATTTTAAACAAAAACCCAGGCTATACGATTTTATGAACGTGAC 669
Qy 1198 TCCTCAATCGTCACATGACAAATGACATTTTCGTTACGATTTTACCAATGATCAAGAG 1257
Db 670 TCCTCAATCGTCACATGACAAATGACATTTTCGTTACGATTTTACCAATGATCAAGAG 729
Qy 1258 TTTACTTACCGTGTAAAAATCGGGAAACAGCTTATAGGATCAATAAAAAATCTGCTCTG 1317
Db 730 TTTACTTACCGTGTAAAAATCGGGAAACAGCTTATAGGATCAATAAAAAATCTGCTCTG 789
Qy 1318 AATGAAGAAATAAACACACACGACCTGTCTCTGAGAAATATTTACGCTTAAAAAAGGG 1377
Db 790 AATGAAGAAATAAACACACACGACCTGTCTCTGAGAAATATTTACGCTTAAAAAAGGG 849
Qy 1378 GAAAAGCGTATGATCCCTTTGATCGAGTCACTTGAAACTGTGTCAACATCAAAATACGTT 1437
Db 850 GAAAAGCGTATGATCCCTTTGATCGAGTCACTTGAAACTGTGTCAACATCAAAATACGTT 909
Qy 1438 GATGTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGTAGCGAACCTAAC 1497

Db 910 GATGTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 969
Qy 1498 TTAGACTTTCAGAGATTTATACGATCCTCGTGATAGGCTAAACTACTCTCAACAACATCTC 1557
Db 970 TTAGACTTTCAGAGATTTATACGATCCTCGTGATAGGCTCAAACTACTCTCAACAACATCTC 1029
Qy 1558 GATGCTTTTGGTATTATGGAATATACCTTAACTGGAAGAGTAGAGGATAATCAGCATGAC 1617
Db 1030 GATGCTTTTGGTATTATGGAATATACCTTAACTGGAAGAGTAGAGGATAATCAGCATGAC 1089
Qy 1618 ACCAACCGTATCAATACCGTTTATATGGCAAGCGACCCGAGGAGAGAATGCTAGCTAT 1677
Db 1090 ACCAACCGTATCAATACCGTTTATATGGCAAGCGACCCGAGGAGAGAATGCTAGCTAT 1149
Qy 1678 CATTTAGCCTCATATAAGATCGTTATACCGAAGAGAACGAGAGATTTTACAGCTACTG 1737
Db 1150 CATTTAGCCTCATATAAGATCGTTATACCGAAGAGAACGAGAGATTTTACAGCTACTG 1209
Qy 1738 CGTTATACGGGACACCTATACCTGTATACCCCTAAACGACAAATAA 1782
Db 1210 CGTTATACGGGACACCTATACCTGTATACCCCTAACGACAAATAA 1254
RESULT 8
LOCUS AR068768 2566 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 19 from patent US 5854049.
ACCESSION AR068768
VERSION AR068768.1 GI:6000975
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2566)
AUTHORS Reed,G.L.
TITLE Plasmid-resistant streptokinase
JOURNAL Patent: US 5854049-A 19-DEC-1998;
FEATURES Location/Qualifiers
source 1. .2566
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 69.0%; Score 1230.4; DB 6; Length 2566;
Best Local Similarity 99.8%; Pred. No. 1.7e-310;
Matches 1243; Conservative 0; Mismatches 1; Indels 2; Gaps 1;
Qy 537 TATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCCAAATTTGGTTGT 596
Db 896 TATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCCAAATTTAGTTGT 955
Qy 597 TAGCGTTGCTGTGCTGTTTGGAGGGAGCAATCAAGACATTTAGTCTTAAATTTTGTGAAT 656
Db 956 TAGCGTTGCTGTGCTGTTTGGAGGGAGCAATCAAGACATTTAGTCTTAAATTTTGTGAAT 1015
Qy 657 CGATCTAAACATACGACCTGCTCATGAGGAAAGACAGACAGCGCTTAAAGTCCAAATC 716
Db 1016 CGATCTAAACATACGACCTGCTCAT - AGGAAAGACAGACAGCGCTTAAAGTCCAAATC 1073
Qy 717 AAAACCAATTTGCTACTGATAGTGGCGGATGTCAATAAACTTGAGAAAGCTGACTTACT 776
Db 1074 AAAACCAATTTGCTACTGATAGTGGCGGATGTCAATAAACTTGAGAAAGCTGACTTACT 1133
Qy 777 AAAGGCTATTTCAGAAACAAATTTGATCGCTAACGTCCTCAAGTAAACGACGACTCTTTGAGGT 836
Db 1134 AAAGGCTATTTCAGAAACAAATTTGATCGCTAACGTCCTCAAGTAAACGACGACTCTTTGAGGT 1193
Qy 837 CATTTGATTTTGAAGCGATGCAACCATTTAATGATCGAAACGGCAAGGTCTACTTTGCTGA 896
Db 1194 CATTTGATTTTGAAGCGATGCAACCATTTAATGATCGAAACGGCAAGGTCTACTTTGCTGA 1253
Qy 897 CAAGATGTTTGGTAACTTGGCCGACCCCAACCTGTCCAAGATTTTGTGCTAACGCGACA 956

|||||
1254 CAAAGATGGTTCGGTAACCTTTGCCGACCCAACTGTCCAAGAAATTTTGTCTAAGCGGACA 1313
QY TGTGCGCGTTAGACCATATAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGT 1016
Db TGTGCGCGTTAGACCATATAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGT 1373
QY GGAATATACCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTACAGACCAAGTCTCAA 1076
Db GGAATATACCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTACAGACCAAGTCTCAA 1433
QY AGATACCTAAGCTATTGAAAAACACATAGCTATCCGTTGACACCAATCAATCTCAAGAAATTA 1493
Db AGATACCTAAGCTATTGAAAAACACATAGCTATCCGTTGACACCAATCAATCTCAAGAAATTA 1493
QY AGCTCAAGCACAAGACATTTTAAACAAAAACCAACCCAGGCTATACGATTTTGAACGCTGA 1196
Db AGCTCAAGCACAAGACATTTTAAACAAAAACCAACCCAGGCTATACGATTTTGAACGCTGA 1553
QY CTCCTCAATCGTCACTCATGACAAATGCAATTTCCGTACGATTTTACCAATGGATCAAGA 1256
Db CTCCTCAATCGTCACTCATGACAAATGCAATTTCCGTACGATTTTACCAATGGATCAAGA 1613
QY GTTTACTTACCGTGTAAATAATCGGAAACAGCTTATAGGATCAATAAAAAATCTGGTCT 1316
Db GTTTACTTACCGTGTAAATAATCGGAAACAGCTTATAGGATCAATAAAAAATCTGGTCT 1673
QY GAAATGAAGAAATAAACAACACTGACTGTCTCTGAGAAATATTACGTCCTTAAAAAAGG 1376
Db GAAATGAAGAAATAAACAACACTGACTGTCTCTGAGAAATATTACGTCCTTAAAAAAGG 1733
QY GGAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAAACCTGTTTACCATCAATAAGT 1436
Db GGAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAAACCTGTTTACCATCAATAAGT 1793
QY TGAATGTCGATACCAACGAAATGCTTAAAGAGTGAGCAGCTCTTAAACAGTACGGAACGTPAA 1496
Db TGAATGTCGATACCAACGAAATGCTTAAAGAGTGAGCAGCTCTTAAACAGTACGGAACGTPAA 1853
QY CTTAGACTTACAGAGATTTATACGATCTCTGATAAGGCTTAACTTACTCTACAACTCT 1556
Db CTTAGACTTACAGAGATTTATACGATCTCTGATAAGGCTTAACTTACTCTACAACTCT 1913
QY CGATGCTTTTGGTATTATGGACTATACCTTAACTGGAAAAGTAGAGGATATACGATGA 1616
Db CGATGCTTTTGGTATTATGGACTATACCTTAACTGGAAAAGTAGAGGATATACGATGA 1973
QY CACCAACCGGTATCATAAACCGTTTATATGGCAAGCGACCCGAAAGGAGAGAAATGCTAGCTA 1676
Db CACCAACCGGTATCATAAACCGTTTATATGGCAAGCGACCCGAAAGGAGAGAAATGCTAGCTA 2033
QY TCATTTAGCCTATGATAAAGATCGTTATACGGAAGAGAACGAGAAAGTTTACAGTACCT 1736
Db TCATTTAGCCTATGATAAAGATCGTTATACGGAAGAGAACGAGAAAGTTTACAGTACCT 2093
QY GCCTTATACAGGACACCTTATACCTGATAACCTTAAACGACAAATAA 1782
Db GCCTTATACAGGACACCTTATACCTGATAACCTTAAACGACAAATAA 2139

RESULT 9
A20027
LOCUS A20027 1458 bp DNA linear PAT 14-JUL-1995
DEFINITION Hirudin-streptokinase fusion protein.
ACCESSION A20027
VERSION A20027.1 GI:1247862
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 1458)
AUTHORS

PROTEINS AND NUCLEIC ACIDS
Patent: WO 9109125-A 35 27-JUN-1991;
Location/Qualifiers
1. .1458
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
ORIGIN
Query Match 68.7%; Score 1223.4; DB 6; Length 1458;
Best Local Similarity 98.3%; Pred. No. 1.1e-308;
Matches 1236; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
QY 526 ACCGATGTTTCGTATTCCTGAGACCTGAGTGGCTCTAGACCGTCCATCTGTCAACAACAGC 585
Db 196 ATCGAAGGTAGAAATTCCTGAGACCTGAGTGGCTCTAGACCGTCCATCTGTCAACAACAGC 255
QY 586 CAATTGTTGTTAGCGTTGCTGTACTGTGTAGGGGACGAATCAAGACATTAAGTCTTAAA 645
Db 256 CAATTGTTGTTAGCGTTGCTGTACTGTGTAGGGGACGAATCAAGACATTAAGTCTTAAA 315
QY 646 TTTTGTGAATCGATCTAACATCAGACCTGCTCATGGAGGAAGACAGAGCAAGGCTTA 705
Db 316 TTTTGTGAATTCGACCTAAACATCAGACCTGCTCATGGAGGAAGACAGAGCAAGGCTTA 375
QY 706 AGTCCAAAATCAAAACCAATTTGCTACTGATAGTGGCGCGATGTCACATAAACTTTGAGAAA 765
Db 376 AGTCCAAAATCAAAACCAATTTGCTACTGATAGTGGCGCGATGCCATTAATTTGAGAAA 435
QY 766 GCTGACTTACTAAAGGCTATTCAAGAACAAATTTGATCGCTAAAGTCCACAGTAAACGACGAC 825
Db 436 GCTGACTTACTAAAGGCTATTCAAGAACAAATTTGATCGCTAAAGTCCACAGTAAACGACGAC 495
QY 826 TACTTTGAGTCAATTTGTTGCAAGCGATGCAACCAATTAATCTGATCGAAACGGCAAGGTC 885
Db 496 TACTTTGAGTCAATTTGTTGCAAGCGATGCAACCAATTAATCTGATCGAAACGGCAAGGTC 555
QY 886 TACTTTGCTGACAAAGATGTTTCGGTAACTTTCGCGACCCCAACCTGTCGAAGATTTTGG 945
Db 556 TACTTTGCTGACAAAGATGTTTCGGTAACTTTCGCGACCCCAACCTGTCGAAGATTTTGG 615
QY 946 CTAAGCGGACATGTGCGCGTTAGACCATATAAGAAAAAAACCAATAAACAACCAAGCGAAA 1005
Db 616 CTAAGCGGACATGTGCGCGTTAGACCATATAAGAAAAAAACCAATAAACAACCAAGCGAAA 675
QY 1006 TCTGTTGATGTGGAATATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCA 1065
Db 676 TCTGTTGATGTGGAATATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCA 735
QY 1066 CCAGGCTCTCAAGATACTTAAGCTATTGAAAACCACTAGCTATCGGTGACACCATCATCT 1125
Db 736 CCAGGCTCTCAAGATACTTAAGCTATTGAAAACCACTAGCTATCGGTGACACCATCATCT 795
QY 1126 CAAGAATTTACTAGCTCAAGCACAAGACATTTTAAACAAAAAACCAACCAAGGCTATACGATT 1185
Db 796 CAAGAATTTACTAGCTCAAGCACAAGACATTTTAAACAAAAAACCAACCAAGGCTATACGATT 855
QY 1186 TATGAACGCTGACTCCTCAATCGTCACTCATGCAATTTTCCGTACGATTTTATACCA 1245
Db 856 TATGAACGCTGACTCCTCAATCGTCACTCATGCAATTTTCCGTACGATTTTATACCA 915
QY 1246 ATGGATCAAGAGTTTACTTTACCGTGTAAAAATTCGGGAACAAGCTTTATAGGATCAATAAA 1305
Db 916 ATGGATCAAGAGTTTACTTTACCATGTCAAAAAATTCGGGAACAAGCTTTATAGGATCAATAAA 975
QY 1306 AAATCTGCTCTGAATGAAGAAATAAACAACCTGACCTGATCTCTGAGAAATATTACGTC 1365
Db 976 AAATCTGCTCTGAATGAAGAAATAAACAACCTGACCTGATCTCTGAGAAATATTACGTC 1035
QY 1366 CTTAAAAAAGGGGAAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTTGAAAACCTGTTACC 1425
Db 1036 CTTAAAAAAGGGGAAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTTGAAAACCTGTTACC 1095

QY 1426 ATCAATACGTTGATGTCGATACCAACGAATGCTAAAGAGTGCAGCAGCTCTTAAACAGCT 1485
 Db 1096 ATCAATACGTTGATGTCGATACCAACGAATGCTAAAGAGTGCAGCAGCTCTTAAACAGCT 1155
 QY 1486 AGCGAAGCTAACTTAGACTTCAGAGATTTATACGATCCTCGTGATAGGCTAAACTACTC 1545
 Db 1156 AGCGAAGCTAACTTAGACTTCAGAGATTTATACGATCCTCGTGATAGGCTAAACTACTC 1215
 QY 1546 TACAACAATCTCGATGCTTTTGGTATTTATGAGCTATACCTTTAACTGCGAAGTAGAGAT 1605
 Db 1216 TACAACAATCTCGATGCTTTTGGTATTTATGAGCTATACCTTTAACTGCGAAGTAGAGAT 1275
 QY 1606 AATACGATGACACCAACCGTATCATAAACCGTTTATATGGCGAAGCACCAGGAGAGAG 1665
 Db 1276 AATACGATGACACCAACCGTATCATAAACCGTTTATATGGCGAAGCACCAGGAGAGAG 1335
 QY 1666 AATGCTAGCTATCAATTTAGCTTAGATTAAGATGCTTTATACCGAAGAACGAGAAGTT 1725
 Db 1336 AATGCTAGCTATCAATTTAGCTTAGATTAAGATGCTTTATACCGAAGAACGAGAAGTT 1395
 QY 1726 TACAGCTACCTGGTTATACAGGACACCTATACCTGATACCTTAACGACAATAA 1782
 Db 1396 TACAGCTACCTGGTTATACAGGACACCTATACCTGATACCTTAACGACAATAA 1452

RESULT 10
 LOCUS I13215 1458 bp DNA linear PAT 26-JUL-1995
 DEFINITION Sequence 42 from patent US 5434073.
 ACCESSION I13215
 VERSION I13215.1 GI:910563
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1458)
 AUTHORS Dawson, K., Hunter, M.G. and Czaplowski, L.G.
 TITLE Fibrinolytic and anti-thrombotic cleavable dimers
 JOURNAL Patent: US 5434073-A 42 18-JUL-1995;
 FEATURES
 source
 1. .1458
 /organism="unknown"
 /mol_type="unassigned DNA"

ORIGIN
 Query Match 68.7%; Score 1223.4; DB 6; Length 1458;
 Best Local Similarity 98.3%; Pred. No. 1.1e-308;
 Matches 1236; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
 QY 526 ACCGATGTTGTTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGC 585
 Db 196 ATCGAAGGTAGAAATTTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGC 255
 QY 586 CAAFTGGTTGTTAGCGTGTCTGCTAGTGGGAGGAGCAATCAAGACATTTAGTCTTAAA 645
 Db 256 CAAFTGGTTGTTAGCGTGTCTGCTAGTGGGAGGAGCAATCAAGACATTTAGTCTTAAA 315
 QY 646 TTTTGGAAATCGATCTAAATCAACGACCTGCTCATGAGGAGGAGGAGGAGGAGGCTTA 705
 Db 316 TTTTGGAAATCGATCTAAATCAACGACCTGCTCATGAGGAGGAGGAGGAGGAGGCTTA 375
 QY 706 AGTCCAAATCAAAACCAATTTGCTAGTGGGCGGATGTCACATTAACCTTGAGAAA 765
 Db 376 AGTCCAAATCAAAACCAATTTGCTAGTGGGCGGATGTCACATTAACCTTGAGAAA 435
 QY 766 GCTGACTTAAAGGCTATTCAAGAACCAATTCATCGCTAAACCTCCACATGACGAGCAG 825
 Db 436 GCTGACTTAAAGGCTATTCAAGAACCAATTCATCGCTAAACCTCCACATGACGAGCAG 495
 QY 826 TACTTTGAGGTCAATGATTTTGAAGCGATGCAACCACTTACTGATCGAAACGGCAGGTC 885
 Db 496 TACTTTGAGGTCAATGATTTTGAAGCGATGCAACCACTTACTGATCGAAACGGCAGGTC 555

QY 886 TACTTTGCTGCAAAAGATGGTTGCGTAAACCTTGCAGCACCACCTGTCGAAGAAATTTTGG 945
 Db 556 TACTTTGCTGCAAAAGATGGTTGCGTAAACCTTGCAGCACCACCTGTCGAAGAAATTTTGG 615
 QY 946 CTAAGCGGACATGTCGCGGTTAGACCATATAAAGAAACCAATACAAAACCAAGCGGAA 1005
 Db 616 CTAAGCGGACATGTCGCGGTTAGACCATATAAAGAAACCAATACAAAACCAAGCGGAA 675
 QY 1006 TCTGTTGATGTTGGAATATCTGTACAGTTTACTCCCTTAAACCTGATGAGATTTACGA 1065
 Db 676 TCTGTTGATGTTGGAATATCTGTACAGTTTACTCCCTTAAACCTGATGAGATTTACGA 735
 QY 1066 CCAGGTCTCAAGATATCAAGCTATTGAAACACATAGCTATCGGTGACACCATCACATCT 1125
 Db 736 CCAGGTCTCAAGATATCAAGCTATTGAAACACATAGCTATCGGTGACACCATCACATCT 795
 QY 1126 CAAGAATTAATAGCTCAAGCAACAAGCATTTTAAACAAAACCAACCGGCTATACGAT 1185
 Db 796 CAAGAATTAATAGCTCAAGCAACAAGCATTTTAAACAAAACCAACCGGCTATACGAT 855
 QY 1186 TATGAAGCTGACTCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCA 1245
 Db 856 TATGAAGCTGACTCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCA 915
 QY 1246 ATGATCAAGAGTTTACTTACCGGTTTAAAAATCGGAAACAAAGCTTATAGGATCAATAA 1305
 Db 916 ATGATCAAGAGTTTACTTACCATGTCAAAATCGGGAACAGGTTATGATCANATAA 975
 QY 1306 AAATCTGCTGTAATGAAGAAATAAACAACATGACCTGATCTCTGAGAAATATTTACGTC 1365
 Db 976 AAATCTGCTGTAATGAAGAAATAAACAACATGACCTGATCTCTGAGAAATATTTACGTC 1035
 QY 1366 CTTAAAAAGGGGAAAAGCCGTATGATCCCTTTGATCGAGTCACTTGAACCTGTCACC 1425
 Db 1036 CTTAAAAAGGGGAAAAGCCGTATGATCCCTTTGATCGAGTCACTTGAACCTGTCACC 1095
 QY 1426 ATCAATACCTGATGTCGATACCAACGAATTTGCTAAAAAGTGCAGGAGCTCTTAAACAGCT 1485
 Db 1096 ATCAATACCTGATGTCGATACCAACGAATTTGCTAAAAAGTGCAGGAGCTCTTAAACAGCT 1155
 QY 1486 AGCGAAGCTAACTTTAGACTTTTACAGATTTTATACGATCCTCGTGATAAGGCTTAACTACTC 1545
 Db 1156 AGCGAAGCTAACTTTAGACTTTTACAGATTTTATACGATCCTCGTGATAAGGCTTAACTACTC 1215
 QY 1546 TACAACAATCTCGATGCTTTTGGTATTTATGAGCTATACCTTTAACTGGAAGAGTAGAGAT 1605
 Db 1216 TACAACAATCTCGATGCTTTTGGTATTTATGAGCTATACCTTTAACTGGAAGAGTAGAGAT 1275
 QY 1606 AATCAGATGACACCAACCGTATCATAAACCGTTTATATGGCGAAGCACCAGGAGAG 1665
 Db 1276 AATCAGATGACACCAACCGTATCATAAACCGTTTATATGGCGAAGCACCAGGAGAG 1335
 QY 1666 AATGCTAGCTATCAATTTAGCTTAGATTAAGATGCTTTATACCGAAGAACGAGAAGTT 1725
 Db 1336 AATGCTAGCTATCAATTTAGCTTAGATTAAGATGCTTTATACCGAAGAACGAGAAGTT 1395
 QY 1726 TACAGCTACCTGGTTATACAGGACACCTTATACCTGATACCTTAACGACAATAA 1782
 Db 1396 TACAGCTACCTGGTTATACAGGACACCTTATACCTGATACCTTAACGACAATAA 1452

RESULT 11
 S46536
 LOCUS
 DEFINITION SKC-2=streptokinase 1245 bp DNA linear BCT 08-MAY-1993
 Genomic, 1245 nt].
 ACCESSION S46536
 VERSION S46536.1 GI:257196
 KEYWORDS
 SOURCE Streptococcus dysgalactiae subsp. equisimilis
 ORGANISM Streptococcus dysgalactiae subsp. equisimilis
 Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 Streptococcus.

REFERENCE 1 (bases 1 to 1245)
AUTHORS Estrada,M.P., Hernandez,L., Perez,A., Rodriguez,P., Serrano,R.,
Rubiera,R., Pedraza,A., Padron,G., Antuch,W., de la Fuente,J. et
al.
TITLE High level expression of streptokinase in *Escherichia coli*
JOURNAL Biotechnology (N.Y.) 10 (10), 1138-1142 (1992)
PUBMED 1368792
REMARK GenBank staff at the National Library of Medicine created this
entry [NCBI gisbbsq 115306] from the original journal article.
FEATURES
source
1. .1245
/organism="Streptococcus dysgalactiae subsp. equisimilis"
/mol_type="genomic DNA"
/db_xref="taxon:119602"
/notes="group: C"
1. .1245
/genes="SKC-2"
1. .1245
/gene="SKC-2"
/notes="streptokinase; Mature protein lacks initial Met"
/codon_start=1
/product="SKC-2"
/protein_id="AAC60418.1"
/db_xref="GI:257197"
/translation="IAGPEWLLDRPSVNSQLVSVAGTVEGTNQDISLKFPEIDLTS
RPAHGKTEGLSPKSPKFPATDSGMPHKLKADLLKAIQELIANVHSNDYPEVID
FASDATITDRNGKVPKADGQSVTLPTQVQEPFLSGHVRVPYKEKPIQNOAKSVDV
EXTVOPTPLNDDPRPLKDLTKLLKTLAIGDRTISOLLAQAGSILNKHPRGYIE
ROSSIVTHNDI,LPRTILPMDQETVHVNRQAYEINKSLGNEINNDLISIKYYV
LKGEKPPDFDRSHUKLFTIKYVDVNTNLSKSEQLTASERNLDFRDLYDPRDKAK
LYNNLDAGFMIDYITLTKGVKEDNDHDTNRIITVYMKRPEGENASYHLAYDKORYTEE
EREVSYLRYTGTPIPDNPDK"
ORIGIN
Query Match 58.6%; Score 1222.6; DB 1; Length 1245;
Best Local Similarity 98.9%; Pred. No. 1.8e-308;
Matches 1231; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 538 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCAGCAATGCTGTGTT 597
DB 1 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCAGCAATGCTGTGTT 60
QY 598 AGCGTTGCTGTTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC 657
DB 61 AGCGTTGCTGTTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC 120
QY 658 GATCTAACTACGACCTGCTCTATGGAGGAAAGACAGACAGAGCTTAAAGTCCAAATCA 717
DB 121 GACCTAACATCAGACCTGCTCATGGAGGAAAGACAGACAGAGCTTAAAGTCCAAATCA 180
QY 718 AAACCATTTGCTACTGATAGTGGCGGATGTCATATAAATTGAGAAGCTGACTTACTA 777
DB 181 AAACCATTTGCTACTGATAGTGGCGGATGTCATATAAATTGAGAAGCTGACTTACTA 240
QY 778 AAGGCTATTCAAGAACAAATTGATCGCTTAACCTCCACAGTAAACGACGACTACTTTGAGTC 837
DB 241 AAGGCTATTCAAGAACAAATTGATCGCTTAACCTCCACAGTAAACGACGACTACTTTGAGTC 300
QY 838 ATGATTTTGAAGCGATGCAACCATTTACTGATCGAAACGCAAGGCTTACTTTGCTGAC 897
DB 301 ATGATTTTGAAGCGATGCAACCATTTACTGATCGAAACGCAAGGCTTACTTTGCTGAC 360
QY 898 AAAGATGGTTCGGTACCTTGGCGACCAACCTGTCACAGAAATTTTGGCTTAAGCGGACAT 957
DB 361 AAAGATGGTTCGGTACCTTGGCGACCAACCTGTCACAGAAATTTTGGCTTAAGCGGACAT 420
QY 958 GTGCGGTTAGACCATATAAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATG 1017
DB 421 GTGCGGTTAGACCATATAAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATG 480
QY 1018 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGTATGACGATTTTCAGACAGGCTCTCAA 1077
DB 481 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGTATGACGATTTTCAGACAGGCTCTCAA 540

QY 1078 GATACTAAGCTATTGAAACACCTAGCTATCGGTGACACCATCACATCTCAAGAATTACTA 1137
DB 541 GATACTAAGCTATTGAAACACCTAGCTATCGGTGACACCATCACATCTCAAGAATTACTA 600
QY 1138 GCTCAAGCACAAGCAATTTTAAACAAACACCCAGGCTATACGATTATGAACGTGAC 1197
DB 601 GCTCAAGCACAAGCAATTTTAAACAAACACCCAGGCTATACGATTATGAACGTGAC 660
QY 1198 TCCTCAATCGTCACATCATGACAATGACATTTTCCGTACGATTTTACCAGTCAAGAG 1257
DB 661 TCCTCAATCGTCACATCATGACAATGACATTTTCCGTACGATTTTACCAGTCAAGAG 720
QY 1258 TTCTACTACCGTGTAAACAAACGGAACCTTATAGGATCAATAAAAAAATCTCGTCTG 1317
DB 721 TTCTACTACCGTGTAAACAAACGGAACCTTATAGGATCAATAAAAAAATCTCGTCTG 780
QY 1318 AATGAAGAAATAAACAACACCTGATCTCTGAGAAATATTACGCTTAAAAAAGG 1377
DB 781 AATGAAGAAATAAACAACACCTGATCTCTGAGAAATATTACGCTTAAAAAAGG 840
QY 1378 GAAAGCCGTATGATCCCTTTTGTATCGCAGTCACTTTGAAACTGTTCAACATCAATACGTT 1437
DB 841 GAAAGCCGTATGATCCCTTTTGTATCGCAGTCACTTTGAAACTGTTCAACATCAATACGTT 900
QY 1438 GATGTCGATACCAACGAATTTGCTAAACAGTGAAGCTCTTAAACAGTAGAGGATAATCAGTAA 1497
DB 901 GATGTCGATACCAACGAATTTGCTAAACAGTGAAGCTCTTAAACAGTAGAGGATAATCAGTAA 960
QY 1498 TTAGACTTCAGAGATTTTATACGATCTCTGATAGAGGCTTAAACTCTCTACACCAATCTC 1557
DB 961 TTAGACTTCAGAGATTTTATACGATCTCTGATAGAGGCTTAAACTCTCTACACCAATCTC 1020
QY 1558 GATGCTTTGTTATATCGACTATACCTTAACCTGAAAGTAGAGGATAATCAGTAA 1617
DB 1021 GATGCTTTGTTATATCGACTATACCTTAACCTGAAAGTAGAGGATAATCAGTAA 1080
QY 1618 ACCAACCGTATCATAAACCGTTTATATGGCAAGCCGACCGAGGAGAGAGTCTAGCTAT 1677
DB 1081 ACCAACCGTATCATAAACCGTTTATATGGCAAGCCGACCGAGGAGAGAGTCTAGCTAT 1140
QY 1678 CATTTAGCTATGATAAAGATCGTTTATACCGAAGAACGAGAAAGTTTACAGCTACCTG 1737
DB 1141 CATTTAGCTATGATAAAGATCGTTTATACCGAAGAACGAGAAAGTTTACAGCTACCTG 1200
QY 1738 CGTTATACAGGACACCTTATACCTGATAACCTTACGACCAATAA 1782
DB 1201 CGTTATACAGGACACCTTATACCTGATAACCTTACGACCAATAA 1245
RESULT 12
QY97820 7057 bp DNA linear PAT 20-APR-2004
LOCUS Sequence 26 from Patent WO2004029256.
DEFINITION
ACCESSION QY97820
VERSION QY97820.1 GI:46426093
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1 other sequences; artificial sequences.
AUTHORS selman-houein Sosa,G., aguiar Cabeza,E., gonz lez Quintero,A.D.
and ramos gonz lez,O.
TITLE Vector for the production of transplastomic angiosperm plants
JOURNAL Patent: WO 2004029256-A 26 08-APR-2004;
CENTRO DE INGENIERIA GENETICA Y BIOTECNOLOGIA (CU)
FEATURES
source
1. .7057
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Description of Artificial Sequence: Nucleotide
sequence of DNA fragment from the vector pVTPA-Estrep

between the rice atpB and tobacco rbcL borders."

ORIGIN

Query Match 68.6%; Score 1222.6; DB 6; Length 7057;
 Best Local Similarity 98.9%; Pred. No. 1.9e-308;
 Matches 1231; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 538 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCCAAATGCTGTT 597
 DB 2023 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCCAAATGCTGTT 2082

QY 598 AGCGTTGCTGTTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAATATTTTGAATC 657
 DB 2083 AGCGTTGCTGTTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAATATTTTGAATC 2142

QY 658 GATCTAACATCAGCACTGCTCATGAGGAGAAAGACAGAGCAAGCTTAAGTCCAAATCA 717
 DB 2143 GACCTAACATCAGCACTGCTCATGAGGAGAAAGACAGAGCAAGCTTAAGTCCAAATCA 2202

QY 718 AAACCAATTTGCTACTGATAGTGGCGGATGCTACATAAATCTTGAGAAAGCTGACTACTA 777
 DB 2203 AAACCAATTTGCTACTGATAGTGGCGGATGCTACATAAATCTTGAGAAAGCTGACTACTA 2262

QY 778 AAGGCTATTCAAGACAAATTTGATCGCTTAAGTCCACAGTAACGACGACTACTTTGAGGTC 837
 DB 2263 AAGGCTATTCAAGACAAATTTGATCGCTTAAGTCCACAGTAACGACGACTACTTTGAGGTC 2322

QY 838 ATTGATTTTGAAGCGATGCAACCATTAATGATGCGAAACCGGACGAGTCTACTTTGCTGAC 897
 DB 2323 ATTGATTTTGAAGCGATGCAACCATTAATGATGCGAAACCGGACGAGTCTACTTTGCTGAC 2382

QY 898 AAGATGTTTCGGTACCTTGGCGACCCACCTGCTCCAGAAATTTTGTAAAGCGGCAT 957
 DB 2383 AAGATGTTTCGGTACCTTGGCGACCCACCTGCTCCAGAAATTTTGTAAAGCGGCAT 2442

QY 958 GTGCGGTTAGACCATATAAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 1017
 DB 2443 GTGCGGTTAGACCATATAAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 2502

QY 1018 GAATATACCTGATGATTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGTCCTCAA 1077
 DB 2503 GAATATACCTGATGATTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGTCCTCAA 2562

QY 1078 GATCTAAGCTATTGAAGACACTAGCTATCGGTGACACCATCATCTCAAGAAATTACTA 1137
 DB 2563 GATCTAAGCTATTGAAGACACTAGCTATCGGTGACACCATCATCTCAAGAAATTACTA 2622

QY 1138 GCTCAAGCACAAGCAATTTTAAACAAAAACCCAGGCTATACGATTTTATGAACGTGAC 1197
 DB 2623 GCTCAAGCACAAGCAATTTTAAACAAAAACCCAGGCTATACGATTTTATGAACGTGAC 2682

QY 1198 TCCTCAATCGTCACTCATGACATGACATTTTCCGTCGATCGATTTTACCAATGGATCAAGAG 1257
 DB 2683 TCCTCAATCGTCACTCATGACATGACATTTTCCGTCGATCGATTTTACCAATGGATCAAGAG 2742

QY 1258 TTTACTACCGTGTAAATAATCGGGAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 1317
 DB 2743 TTTACTACCGTGTAAATAATCGGGAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 2802

QY 1318 AATGAAGAAATAACCAACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAAGGG 1377
 DB 2803 AATGAAGAAATAACCAACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAAGGG 2862

QY 1378 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAAATCTGTTCCACCAATCAATACGTT 1437
 DB 2863 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAAATCTGTTCCACCAATCAATACGTT 2922

QY 1438 GATGTCGATACCAACGAATGCTTAAAGATGAGAGAGCTCTTAAAGCTAGCGAAACGTAAC 1497
 DB 2923 GATGTCGATACCAACGAATGCTTAAAGATGAGAGAGCTCTTAAAGCTAGCGAAACGTAAC 2982

QY 1498 TTAGACTTCAGAGATTTTACAGTCTCTGATAGAGCTTAACTTCTTACACAACTC 1557

2983 TTAGACTTCAGAGATTTTATACGATCTCTGATAGGCTAAACTACTCTACAACAATCTC 3042

QY 1558 GATGCTTTTGGTATTATGACTATACCTTAACCTGGAAGTAGAGGATAATACAGATGAC 1617
 DB 3043 GATGCTTTTGGTATTATGACTATACCTTAACCTGGAAGTAGAGGATAATACAGATGAC 3102

QY 1618 ACCAACCGTATCATAAACCGTTTATATGGCAAGCGACCCGAAGGAGAGAATGCTAGCTAT 1677
 DB 3103 ACCAACCGTATCATAAACCGTTTATATGGCAAGCGACCCGAAGGAGAGAATGCTAGCTAT 3162

QY 1678 CATTTAGCCTATGATAAAGATCGTTTATACCGAAGAACGAGAAGTTTACAGCTACCTG 1737
 DB 3163 CATTTAGCCTATGATAAAGATCGTTTATACCGAAGAACGAGAAGTTTACAGCTACCTG 3222

QY 1738 CGTTTATACAGGACACCTATACCTGATTAACCTTAACGACAATAA 1782
 DB 3223 CGTTTATACAGGACACCTATACCTGATTAACCTTAACGACAATAA 3267

RESULT 13

A20006

LOCUS

DEFINITION

SEQ ID NO: 14; Streptokinase gene.

ACCESSION

A20006

VERSION

A20006.1 GI:1247579

KEYWORDS

SOURCE

ORGANISM

Streptococcus dysgalactiae subsp. equisimilis

Streptococcus dysgalactiae subsp. equisimilis

Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

REFERENCE

1 (bases 1 to 1335)

AUTHORS

PROTEINS AND NUCLEIC ACIDS

TITLE

Patent: WO 9109125-A 14 27-JUN-1991;

JOURNAL

Location/Qualifiers

FEATURES

source

1. 1335

/organism="Streptococcus dysgalactiae subsp. equisimilis"

/mol_type="unassigned DNA"

/sub_species="equisimilis"

/db_xref="taxon:119602"

7. 1329

/codon_start=1

/transl_table=1

/product="streptokinase"

/protein_id="CAA01484.1"

/db_xref="GI:1247580"

/translation="MKNYLSFGMPALLFALTFTGTVNSVQAIAGPEWLLDRPSVNSQL

VVSVAGTVEGTNODISLKPFEDITSRPAHGKTEQGLSPKSPKPDATDSGAMPKLEK

ADLLKALQEQOLIANVHSDYFEVIDPASDATITDRNGKVPADKDSVTLPTQVQS

FLLSGHVRVRYPEKPIQONAKSVDSVETVOFTPLNPDDDFRPGIKDKTLKLLAIG

TITSQELLAQAQSIILNKTHPGYTIYERDSSIVTHDNDI FRTILPMDQEFYVKNRQ

AYEINKSGLINEINNTDLIS EKYVILKGEKPYDFRSHLKLFTI KYVDVNTNELL

KSEOLLTASERNLDFRLYDPRDKALLNNLDAFIMDYTLTGKVEDNDDHTNRIIT

VYMKRPEGENASVHLAYDKDRYTEBEREVSYLRYTGTPTPDNPDK"

CDS

Query Match 68.5%; Score 1220.4; DB 6; Length 1335;
 Best Local Similarity 98.7%; Pred. No. 6.8e-308;
 Matches 1230; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 537 TATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAAATGTTGT 596
 DB 84 TATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAAATGTTGT 143

QY 597 TAGCGTTGCTGTTAGTGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAAT 656
 DB 144 TAGCGTTGCTGTTAGTGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAAT 203

QY 657 CGATCTAACATCAGACCTGCTCATGGAGGAAAGACAGAGCAAGCTTTAAGTCCAAATC 716
 DB 204 TGACCTAACATCAGACCTGCTCATGGAGGAAAGACAGAGCAAGCTTTAAGTCCAAATC 263

QY 717 AAACCAATTTGCTACTGATAGTGGCGGATGTCATATAAATTGAGAAAGCTGACTTACT 776

Db 264 AAAACCAATTGGCTACTGATAGTGGCGATGCCACATAAACTTGAAAAAGCTGACTTACT 323
Qy 777 AAAGGCTATTCAAGAAACAATTGATCGCTAACGTCACAGTAACGACGACTACTTTTGAGGT 836
Db 324 AAAGGCTATTCAAGAACAAATTGATCGCTACGTCACAGTAACGACGACTACTTTTGAGGT 383
Qy 837 CATTGATTTTGGACGCGATGAACCAATTAATGATCGAAACGCGAAGGCTACTTTTGCTGA 896
Db 384 CATTGATTTTGGACGCGATGAACCAATTAATGATCGAAACGCGAAGGCTACTTTTGCTGA 443
Qy 897 CAAAGATGTTCCGTTACCTTGGCGACCGACCTGTCACAGAAATTTTGGTAAAGCGGACA 956
Db 444 CAAAGATGTTCCGTTACCTTGGCGACCGACCTGTCACAGAAATTTTGGTAAAGCGGACA 503
Qy 957 TGTGCGGTTTGAACCAATATAAGAAAAACCAATACAAAAACCAAGCAAAATCTGTTGATGT 1016
Db 504 TGTGCGGTTTGAACCAATATAAGAAAAACCAATACAAAAACCAAGCAAAATCTGTTGATGT 563
Qy 1017 GGAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCAAGTCTCAA 1076
Db 564 GGAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCAAGTCTCAA 623
Qy 1077 AGATCTAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTACT 1136
Db 624 AGATCTAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTACT 683
Qy 1137 AGCTCAAGCAACAAGCAATTTTAAACAAAAACCCAGGCTATACGATTTATGAACGTGA 1196
Db 684 AGCTCAAGCAACAAGCAATTTTAAACAAAAACCCATCCAGGCTATACGATTTATGAACGTGA 743
Qy 1197 CTCCTCAATCGTCACTCATGACAAATGATTTTCCGTACGATTTTACCAATGGAATCAAGA 1256
Db 744 CTCCTCAATCGTCACTCATGACAAATGATTTTCCGTACGATTTTACCAATGGAATCAAGA 803
Qy 1257 GTTTACTTTACCGTGTAAATAATCGGAAACAAAGCTTATAGGATCAATAAAAAATCTGGTCT 1316
Db 804 GTTTACTTTACCAATGTCAAAAAATCGGAAACAAAGCTTATAGGATCAATAAAAAATCTGGTCT 863
Qy 1317 GAATGAAGAAATAAACAACGACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGG 1376
Db 864 GAATGAAGAAATAAACAACGACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGG 923
Qy 1377 GGAAGAACCGGTATGATCCCTTTGATCGCAGTCACTTGAAACCTGTTTACCACATCAATAGCT 1436
Db 924 GGAAGAACCGGTATGATCCCTTTGATCGCAGTCACTTGAAACCTGTTTACCACATCAATAGCT 983
Qy 1437 TGATGTCGATACCAACGAAATTGCTTAAAGAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAA 1496
Db 984 TGATGTCGATACCAACGAAATTGCTTAAAGAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAA 1043
Qy 1497 CTTAGACTTCAGAGATTTATACGATCCCTGATGTAAGGCTTAACTCTCTACAAATCT 1556
Db 1044 CTTAGACTTCAGAGATTTATACGATCCCTGATGTAAGGCTTAACTCTCTACAAATCT 1103
Qy 1557 CGATGCTTTTGGTATTATGACATATACCTTAACTGGAAGAGTACGATTAATACGATGA 1616
Db 1104 CGATGCTTTTGGTATTATGACATATACCTTAACTGGAAGAGTACGATTAATACGATGA 1163
Qy 1617 CACCAACCGGTATCATAAACCGTTTATATGGGCAAGCGACCCGAGGAGAGATGCTAGCTA 1676
Db 1164 CACCAACCGGTATCATAAACCGTTTATATGGGCAAGCGACCCGAGGAGAGATGCTAGCTA 1223
Qy 1677 TCATTTAGCTATGATAAGATCGTTTATACGGAAGAAACGAGAGTTTACAGTACCT 1736
Db 1224 TCATTTAGCTATGATAAGATCGTTTATACGGAAGAAACGAGAGTTTACAGTACCT 1283
Qy 1737 GCGTTATACGAGACACCTTATACCTGATTAACCTTAACCGCAAAATAA 1782
Db 1284 GCGTTATACGAGACACCTTATACCTGATTAACCTTAACCGCAAAATAA 1329

I13194
LOCUS I13194 1335 bp DNA linear PAT 26-JUL-1995
DEFINITION Sequence 14 from patent US 5434073.
ACCESSION I13194
VERSION I13194.1 GI:910542
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1335)
AUTHORS Dawson,K., Hunter,M.G. and Czaplewski,L.G.
TITLE Fibrinolytic and anti-thrombotic cleavable dimers
JOURNAL Patent: US 5434073-A 14 18-JUL-1995;
FEATURES Location/Qualifiers
1..1335
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 68.5%; Score 1220.4; DB 6; Length 1335;
Best Local Similarity 98.7%; Pred. No. 6.8e-308;
Matches 1230; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
Qy 537 TATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAAACAACAGCCAATTGGTTGT 596
Db 84 TATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAAACAACAGCCAATTAGTTGT 143
Qy 597 TAGCGTTCTGCTACTGTTGGGGACGAATCAAGACATTAGTCTTAATTTTTTGAAT 656
Db 144 TAGCGTTCTGCTACTGTTGGGGACGAATCAAGACATTAGTCTTAATTTTTTGAAT 203
Qy 657 CGATCTAACATCACGACCTGCTCATGGAGGAAGACAGACGAGCTTAAAGTCCAAAATC 716
Db 204 TGACCTAACATCACGACCTGCTCATGGAGGAAGACAGACGAGCTTAAAGTCCAAAATC 263
Qy 717 AAAACCAATTGCTACTGATAGTGGCGCGATGTGCATATAAACTTTGAGAAGCTGACTTACT 776
Db 264 AAAACCAATTGCTACTGATAGTGGCGCGATGCCACATAAACTTTGAGAAGCTGACTTACT 323
Qy 777 AAAGGCTATTCAAGAAACAATTGATCGCTAACGTCACAGTAACGACGACTACTTTGAGGT 836
Db 324 AAAGGCTATTCAAGAAACAATTGATCGCTAACGTCACAGTAACGACGACTACTTTGAGGT 383
Qy 837 CATTGATTTTTCAGACGATGCAACCAATTAATGATCGAAACGCGAAGGCTACTTTTGCTGA 896
Db 384 CATTGATTTTTCAGACGATGCAACCAATTAATGATCGAAACGCGAAGGCTACTTTTGCTGA 443
Qy 897 CAAAGATGTTCCGTTACCTTGGCGACCGACCTGTCACAGAAATTTTGGTAAAGCGGACA 956
Db 444 CAAAGATGTTCCGTTACCTTGGCGACCGACCTGTCACAGAAATTTTGGTAAAGCGGACA 503
Qy 957 TGTGCGGTTTGAACCAATATAAGAAAAACCAATACAAAAACCAAGCAAAATCTGTTGATGT 1016
Db 504 TGTGCGGTTTGAACCAATATAAGAAAAACCAATACAAAAACCAAGCAAAATCTGTTGATGT 563
Qy 1017 GGAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCAAGTCTCAA 1076
Db 564 GGAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCAAGTCTCAA 623
Qy 1077 AGATCTAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTACT 1136
Db 624 AGATCTAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTACT 683
Qy 1137 AGCTCAAGCAACAAGCAATTTTAAACAAAAACCCAGGCTATACGATTTATGAACGTGA 1196
Db 684 AGCTCAAGCAACAAGCAATTTTAAACAAAAACCCATCCAGGCTATACGATTTATGAACGTGA 743
Qy 1197 CTCCTCAATCGTCACTCATGACAAATGATTTTCCGTACGATTTTACCAATGGAATCAAGA 1256
Db 744 CTCCTCAATCGTCACTCATGACAAATGATTTTCCGTACGATTTTACCAATGGAATCAAGA 803
Qy 1257 GTTTACTTTACCGTGTAAATAATCGGAAACAAAGCTTATAGGATCAATAAAAAATCTGGTCT 1316

Db 804 GTTACTTACCATGTCAAAAATCGGAAACAGCTTATGAGATCAATAAAAAATCGTCT 863
 QY 1317 GAATGAAGAAATATAACACACTGACCTGATCTCTGAGAAATATTACGCTCTTAAAAAAGG 1376
 Db 864 GAATGAAGAAATATAACACACTGACCTGATCTCTGAGAAATATTACGCTCTTAAAAAAGG 923
 QY 1377 GGAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCAACCATCAAAATACGT 1436
 Db 924 GGAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCAACCATCAAAATACGT 983
 QY 1437 TGATGTCGATACCAACGAATTTGATAAAGTAGGAGCAGCTCTTAACAGTAGCGAAGCTAA 1496
 Db 984 TGATGTCGATACCAACGAATTTGATAAAGTAGGAGCAGCTCTTAACAGTAGCGAAGCTAA 1043
 QY 1497 CTTAGACTTTCAGAGATTTATACGATCCTCGTGATAGGCTTAACTACTCTACACCAATCT 1556
 Db 1044 CTTAGACTTTCAGAGATTTATACGATCCTCGTGATAGGCTTAACTACTCTACACCAATCT 1103
 QY 1557 CGATGCTTTTGGTATTATGGACTATATACCTTAACTGGAAGAGTAGAGATTAATCACGATGA 1616
 Db 1104 CGATGCTTTTGGTATTATGGACTATATACCTTAACTGGAAGAGTAGAGATTAATCACGATGA 1163
 QY 1617 CACCAACCGTATCATTAACCGTTTATATGGCAAGCGACCCGAGGAGAGATGCTAGCTA 1676
 Db 1164 CACCAACCGTATCATTAACCGTTTATATGGCAAGCGACCCGAGGAGAGATGCTAGCTA 1223
 QY 1677 TCATTTAGCCTATGATAAGATCGTTATACCGAAGAGAGAGAGATTTACAGCTACCT 1736
 Db 1224 TCATTTAGCCTATGATAAGATCGTTATACCGAAGAGAGAGAGATTTACAGCTACCT 1283
 QY 1737 GCGTTATACAGGACACCTTATACCTGATTAACCCCTAACGACAAATAA 1782
 Db 1284 GCGTTATACAGGACACCTTATACCTGATTAACCCCTAACGACAAATAA 1329

RESULT 15
 LOCUS A20016 1512 bp DNA linear PAT 14-JUL-1995
 DEFINITION SEQ ID NO: 24; Nucleotide sequence for streptokinase fused to yeast alpha-factor.
 ACCESSION A20016
 VERSION A20016.1 GI:1247850
 KEYWORDS synthetic construct
 SOURCE synthetic construct
 ORGANISM other sequences; artificial sequences.
 REFERENCE 1 (bases 1 to 1512)
 AUTHORS PROTEINS AND NUCLEIC ACIDS
 TITLE Patent: WO 9109125-A 24 27-JUN-1991;
 JOURNAL Location/Qualifiers
 FEATURES
 source
 1. .1512
 /organism="synthetic construct"
 /mol_type="unassigned DNA"
 /db_xref="taxon:32630"
 7. .1506
 /codon_start=1
 /transl_table=11
 /product="streptokinase fused to yeast alpha-factor"
 /protein_id="CAA01487.1"
 /db_xref="GI:1247851"
 /translation="WRPSPSIFATLVFAASSALAAPVNTTDETAQIPABAVIGYLDL
 EGDVAVLPFNSNTNGLFINTTASIAKEEGVSLDKRIAGPBLWLDPRVNSQ
 LVSVAGTVEGNTQISLKFIDLTSPAHGKTEQGLSPKSPKPTADSGAMPHKLE
 KADLLKATQELIANVHNSNDYFEVIDFASDATITDRNGKVYPADKDGSVTLPTQVQ
 BELLSHVRVYKPKIQONAKSVDEVTQFTPLNPDPPDRPGLKDTKLKLTLAG
 DTITQELLQAQSLNKHTRPGVTIYERDSSIVTHDNDIFRILPMDQEFYHVKNRE
 QAYEINKSGLEENINTDLISEKTYVLKRGKPPDFORSHLKLFTIKYVDVNTNEL
 LKSEQLLTASENLDFRDLYDRDKAKLYNNLDFAFGIMDYTLTGKVEDNDTNRIL
 TVYMKRPEGENASYHLAYDKORYTEEREVSYLYRTGTPIPDNPNDK"

Query Match 68.5%; Score 1220.4; DB 6; Length 1512;

Best Local Similarity 98.3%; Pred. No. 6.8e-308;
 Matches 1233; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
 QY 529 GATGTTGCTATGCTGGACCTGAGTGGCTGTAGACCGTCTCATCTGTCAACACAGCCAA 588
 Db 253 GATAAAAAGAAATGCTGGACCTGAGTGGCTGTAGACCGTCTCATCTGTCAACACAGCCAA 312
 QY 589 TTGGTTGTTAGCGTTGCTGGTACTGTTGAGGGGAGCAATCAAGACATTAAGTCTTAAATTT 648
 Db 313 TTAGTTGTTAGCGTTGCTGGTACTGTTGAGGGGAGCAATCAAGACATTAAGTCTTAAATTT 372
 QY 649 TTTGAAATTCGATCTAAACATACGACCTGCTCATGAGGAAAGACAGAGCAAGGCTTAAAGT 708
 Db 373 TTTGAAATTCGATCTAAACATACGACCTGCTCATGAGGAAAGACAGAGCAAGGCTTAAAGT 432
 QY 709 CAAAAATCAAAACCAATTTGCTACTGATAGTGGCGGATGTACATATAAATTGAGAAAGCT 768
 Db 433 CAAAAATCAAAACCAATTTGCTACTGATAGTGGCGGATGTACATATAAATTGAGAAAGCT 492
 QY 769 GACTTACTTAAAGGCTATTTCAGAAACAATTGATCGCTAACGTCACAGTAAACGACGACTAC 828
 Db 493 GACTTACTTAAAGGCTATTTCAGAAACAATTGATCGCTAACGTCACAGTAAACGACGACTAC 552
 QY 829 TTTGAGGTCATTGATTTTGCAGGCGATGCAACCACTTACTGATCGAAACGCGAAGCTCTAC 888
 Db 553 TTTGAGGTCATTGATTTTGCAGGCGATGCAACCACTTACTGATCGAAACGCGAAGCTCTAC 612
 QY 889 TTTGCTGACAAAAGATGTTGCTGTAACCTTGGCGACCCAACTGTCTCAAGAAATTTTGTCTA 948
 Db 613 TTTGCTGACAAAAGATGTTGCTGTAACCTTGGCGACCCAACTGTCTCAAGAAATTTTGTCTA 672
 QY 949 AGCGGACATGCGCGTTAGACCATATAAGAAACCAATACAAAACCAAGCGGAAATCT 1008
 Db 673 AGCGGACATGCGCGTTAGACCATATAAGAAACCAATACAAAACCAAGCGGAAATCT 732
 QY 1009 GTTGATGTGGAATATACGTACAGTTTACTCCCTTAAACCTGTAGTACGATTTTCAGACCA 1068
 Db 733 GTTGATGTGGAATATACGTACAGTTTACTCCCTTAAACCTGTAGTACGATTTTCAGACCA 792
 QY 1069 GGTCTCAAAAGATATAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCTCAA 1128
 Db 793 GGTCTCAAAAGATATAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCTCAA 852
 QY 1129 GAATTACTAGCTCAAGCACAAAGCATTTTAAACAAAACCCAGCGGTATACGATTTAT 1188
 Db 853 GAATTACTAGCTCAAGCACAAAGCATTTTAAACAAAACCCAGCGGTATACGATTTAT 912
 QY 1189 GAACGTGACTCTCAATCGTCACTCATGACAAATGACATTTTCCGTACCAATTTTACCAATG 1248
 Db 913 GAACGTGACTCTCAATCGTCACTCATGACAAATGACATTTTCCGTACCAATTTTACCAATG 972
 QY 1249 GATCAAGAGTTTACTTACCGTGTGTTAAAAATCGGAAACAGCTTATAGGATCAATAAAAAA 1308
 Db 973 GATCAAGAGTTTACTTACCGTGTGTTAAAAATCGGAAACAGCTTATAGGATCAATAAAAAA 1032
 QY 1309 TCTGCTCTGAATGAAGAAAATAACACACACTGACCTGATCTCTGAGAAATATATACGCTCTT 1368
 Db 1033 TCTGCTCTGAATGAAGAAAATAACACACACTGACCTGATCTCTGAGAAATATATACGCTCTT 1092
 QY 1369 AAAAAAGGGGAAAAAGCGGTATGATCCCTTTCATCGCAGTCACTTGAACACTGTTTACCAATC 1428
 Db 1093 AAAAAAGGGGAAAAAGCGGTATGATCCCTTTCATCGCAGTCACTTGAACACTGTTTACCAATC 1152
 QY 1429 AAATACGTTGATGTGCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGTAGC 1488
 Db 1153 AAATACGTTGATGTGCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGTAGC 1212
 QY 1489 GAACGTAACCTTAGACTTCAGAGATTTATACGATCTCTGATATAGGCTTAAACTACTCTAC 1548
 Db 1213 GAACGTAACCTTAGACTTCAGAGATTTATACGATCTCTGATATAGGCTTAAACTACTCTAC 1272
 QY 1549 AACCAATCTCGATGCTTTTGGTATTATGACATATACCTTAACTGGAAGAGTAGGAGATAT 1608

Db	1273	AACAATCTCGATGCTTTTGGTATTATGGACTATACCTTAACCTGGAAAAGTAGAAGATAAT	1332
Qy	1609	CACGATGACACCAACCGTATCATAAACCGTTTATATGGCAAGCGACCCGAAAGGAGAGAAT	1668
Db	1333	CACGATGACACCAACCGTATCATAAACCGTTTATATGGCAAGCGACCCGAAAGGAGAGAAT	1392
Qy	1669	GCTAGCTATCATTTAGCCTATGATAAAGATCGTTATACCGAAGAAACGAGAGGTTTAC	1728
Db	1393	GCTAGCTATCATTTAGCCTATGATAAAGATCGTTATACCGAAGAAACGAGAGGTTTAC	1452
Qy	1729	AGCTACCTGCGTTATACAGGACACCTATACCTGATACCCCTAACGACAAATAA	1782
Db	1453	AGCTACCTGCGTTATACAGGACACCTATACCTGATACCCCTAACGACAAATAA	1506

Search completed: February 1, 2006, 04:17:48
Job time : 9060.25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2006, 10:44:08 ; Search time 1077.25 Seconds
(without alignments)
11024.772 Million cell updates/sec

Title: US-09-940-235-11
Perfect score: 1782
Sequence: 1 tcgcttcacgttcgctcgcg.....ataacctaacgacaaataa 1782

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 332346308 residues
Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq 21.*
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004as.*
13: Geneseqn2004bs.*
14: Geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1782	100.0	1782	3	Aaa37642 Chimeric
2	1684	94.5	2096	3	Aaa37643 Chimeric
3	1245.8	69.9	2385	2	Aax80497 Streptoki
4	1245	69.9	1377	3	Aaa37622 Streptoki
5	1243.4	69.8	1245	3	Aaa37633 S. equisi
6	1241.8	69.7	2030	2	Aaa37633 S. equisi
7	1240.4	69.6	1242	2	Aax80492 Streptoco
8	1240.4	69.6	1254	6	Aba05546 Streptoki
9	1240.4	69.6	8893	6	Aba05547 Maxadil
10	1238.8	69.5	1242	2	Aax16632 Streptoco
11	1235.6	69.3	1242	2	Aax16633 Streptoco
12	1232.2	69.1	1327	3	Aaa37628 Streptoki
13	1230.4	69.0	2566	2	Aat77778 Coding se
14	1223.4	68.7	1458	2	Aaql2162 Factor xa
15	1222.6	68.6	7057	12	Adm01294 Plasmid p
16	1220.4	68.5	1335	2	Aaql2156 Streptoki
17	1220.4	68.5	1512	2	Aaql2158 Streptoki
18	1220.2	68.5	2589	2	Aaql2160 OmpAL str
19	1219.4	68.4	1245	2	Aaq20665 SKC-2 str

20	1216.4	68.3	1467	2	AAQ12490	Aaql2490 Factor Xa
21	1212.4	68.0	1473	2	AAQ05603	Aaql05603 Streptoki
22	1211.4	68.0	1407	1	AAQ070106	Aaql070106 DNA encod
23	1202.8	67.5	2568	1	AAQ05493	Aaql05493 Sequence
24	1198	67.2	1323	2	AAQ29961	Aaql29961 Vector pS
25	1189.6	66.8	1242	5	AAQ82144	Aaql82144 Mutant st
26	1184	66.4	1209	3	AAQ299249	Aaql299249 DNA encod
27	1181.4	66.3	1245	3	AAQ299250	Aaql299250 DNA encod
28	1150.2	64.5	1661	3	AAQ37637	Aaql37637 Chimeric
29	1134.2	63.6	1541	3	AAQ37644	Aaql37644 Chimeric
30	1093.6	61.4	1122	3	AAQ299251	Aaql299251 DNA encod
31	1093.6	61.4	1158	3	AAQ299252	Aaql299252 DNA encod
32	1081.2	60.7	1323	13	ADR83811	Adr83811 S. pyogen
33	1081	60.7	1119	2	AAQ12159	Aaql12159 Truncated
34	1080.8	60.7	2253	2	AAQ12161	Aaql12161 Met-core
35	1079.6	60.6	1473	2	AAQ05604	Aaql05604 Streptoki
36	1078.2	60.5	1320	6	ABN70192	Abn70192 Streptoco
37	1068	59.9	1068	2	AAX80493	Aax80493 Recombina
38	1063.4	59.7	2208	2	AAX83589	Aax83589 Recombina
39	984.6	55.3	1245	10	ADP48644	Adp48644 Streptoco
40	829	46.5	1262	2	AAQ10230	Aaql0230 Synthetic
41	450	25.3	450	2	AAX80494	Aax80494 Recombina
42	233.6	13.1	4727	4	AAQ09265	Aaql09265 pT7-TACS
43	232.8	13.1	1929	13	ADS17488	Adsl7488 Nucleotid
44	232.8	13.1	1929	13	ADR97657	Adr97657 Human fib
45	232.8	13.1	1929	14	ADW44479	Adw44479 Human fib

ALIGNMENTS

RESULT 1
AAA37642
ID AAA37642 standard; DNA; 1782 BP.

XX AAA37642;
XX AC
XX 15-SEP-2003 (revised)
DT 13-OCT-2000 (first entry)
XX Chimeric SK-FBD coding sequence.

XX Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
KW plasminogen; human; fibrinectin; thrombolytic therapy;
KW cardiovascular disorder; ss.

XX Streptococcus dysgalactiae subsp. equisimilis.
OS Homo sapiens.
OS Chimeric.

XX EPI024192-A2.
XX 02-AUG-2000.

XX 23-DEC-1999; 99EP-00310541.
XX 24-DEC-1998; 98IN-DE003825.

XX (COUL) CSIR COUNCIL SCI IND RES.

XX Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
PI Yadav M;

XX WPI; 2000-516032/47.

XX Hybrid streptokinase-fibrin binding domain polypeptides useful for
PT thrombolytic therapy comprises a streptokinase fused with fibrin binding
PT domains of human fibrinectin.

XX Example 5; Fig 21b; 58pp; English.

XX This sequence represents a chimeric streptokinase-fibrin binding domain
CC (SK-FBD) protein coding sequence. The invention relates to a hybrid

CC plasminogen activator (PA) comprises a polypeptide fusion between
 CC streptokinase (SK), which are capable of plasminogen (PG) activation, and
 CC fibrin binding regions of human fibrinectin, which are from fibrin
 CC binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA possesses the
 CC ability to bind with fibrin independently and also characteristically
 CC retains a PG activation ability which becomes evident only after a
 CC pronounced duration, or lag, after exposure of the PA to a suitable
 CC animal or human PG. The hybrid streptokinase-fibrin binding domain
 CC polypeptides are useful in thrombolytic therapy for various kinds of
 CC cardiovascular disorders. The hybrids have enhanced fibrin selectivity as
 CC well as kinetics of plasminogen activation that are distinct from that of
 CC natural streptokinase in being characterised by a temporary delay, or lag
 CC of several minutes in the natural rate of the catalytic conversion of
 CC plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins
 CC can bind tightly with fibrin in blood clots soon after introduction into
 CC the vascular system without significantly activating the circulating
 CC blood plasminogen to plasmin, thus aiding in the localisation of the
 CC plasminogen activation process to the site of pathological thrombus. This
 CC overcomes systemic plasminogen activation encountered during clinical use
 CC of streptokinase. (Updated on 15-SEP-2003 to standardise OS field)

XX Sequence 1782 BP; 573 A; 407 C; 377 G; 425 T; 0 U; 0 Other;

Query Match 100.0%; Score 1782; DB 3; Length 1782;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1782; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	TCGCTTACGTTCCGTCGCGTATCGGTGATTCATTCGTAAACAGTAAGGCAACCCCGC	60
Db	1	TCGCTTACGTTCCGTCGCGTATCGGTGATTCATTCGTAAACAGTAAGGCAACCCCGC	60
Qy	61	CAGCTACGCGGTCTCAAGACAGGAGCAGCATCGGACCGCGCCAGGACCCCA	120
Db	61	CAGCTACGCGGTCTCAAGACAGGAGCAGCATCGGACCGCGCCAGGACCCCA	120
Qy	121	ACGCTGCGCGGAGATCTCGATCCCGGAAATTAATACGATCTCAATAGGAGACCAAC	180
Db	121	ACGCTGCGCGGAGATCTCGATCCCGGAAATTAATACGATCTCAATAGGAGACCAAC	180
Qy	181	GGTTTCCCTCTAGAAATAATTTTGTAACTTTAAAGAGGAGATATACCATGGTCAAGC	240
Db	181	GGTTTCCCTCTAGAAATAATTTTGTAACTTTAAAGAGGAGATATACCATGGTCAAGC	240
Qy	241	ACACAGATCTTACCATAGCTGAGAGTGTGTTGATCATGCTGCGGACCTTCCTATGT	300
Db	241	ACACAGATCTTACCATAGCTGAGAGTGTGTTGATCATGCTGCGGACCTTCCTATGT	300
Qy	301	GGTCGAGAAACGTGGGAGAGCGAGCGGACGCATCACTTGCATCTTCTAGAAATAGATGC	360
Db	301	GGTCGAGAAACGTGGGAGAGCGAGCGGACGCATCACTTGCATCTTCTAGAAATAGATGC	360
Qy	361	AACGATCAGGACACAAAGGACATCTTATAGAAATGGAGACACCTGGAGCAAGAGGATAAT	420
Db	361	AACGATCAGGACACAAAGGACATCTTATAGAAATGGAGACACCTGGAGCAAGAGGATAAT	420
Qy	421	CGAGAAACCTGCTCAGTGATCTGCA CAGGCAACGGCCGAGGAGTGGAGTGTGAG	480
Db	421	CGAGAAACCTGCTCAGTGATCTGCA CAGGCAACGGCCGAGGAGTGGAGTGTGAG	480
Qy	481	AGGCACACCTCTGTGACAGACACATCGAGCGGATCTGGCCCTTCAACCGATGTTCTGATTT	540
Db	481	AGGCACACCTCTGTGACAGACACATCGAGCGGATCTGGCCCTTCAACCGATGTTCTGATTT	540
Qy	541	GCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAAAGAGGAGTGGTGTAGC	600
Db	541	GCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAAAGAGGAGTGGTGTAGC	600
Qy	601	GTTGCTGTACTGTGAGGAGGAGATCAAGACATTAAGTCTTAAATTTTGGAAATCGAT	660
Db	601	GTTGCTGTACTGTGAGGAGGAGATCAAGACATTAAGTCTTAAATTTTGGAAATCGAT	660
Qy	661	CTAACATCAGGACCTGCTCATGGAGGAGAGCAGAGGAGGCTTAAGTCCAAATCAAAA	720
Db	661	CTAACATCAGGACCTGCTCATGGAGGAGAGCAGAGGAGGCTTAAGTCCAAATCAAAA	720

Db	661	CTAACATCAGGACCTGCTCATGGAGGAGAGCAGAGGAGGCTTAAGTCCAAATCAAAA	720
Qy	721	CCATTTTCTACTGATAGTGGCGGATGTCATATAAATCTTGAGAAAGCTGACTTACTTAAG	780
Db	721	CCATTTTCTACTGATAGTGGCGGATGTCATATAAATCTTGAGAAAGCTGACTTACTTAAG	780
Qy	781	GCTATTCAAGAAACAATTTGATCGGTAACTCCACAGTAACGACGACTACTTTGAGGTCAAT	840
Db	781	GCTATTCAAGAAACAATTTGATCGGTAACTCCACAGTAACGACGACTACTTTGAGGTCAAT	840
Qy	841	GATTTTTCGAAGCGATGCAACATTTACTGATCGAAACGGCAAGGTCTACTTTGCTGACAAA	900
Db	841	GATTTTTCGAAGCGATGCAACATTTACTGATCGAAACGGCAAGGTCTACTTTGCTGACAAA	900
Qy	901	GATGGTTCGTTACCTTGGCGACCCCACTGTCTCAAGAAATTTTGTCTAAGGGACATGTG	960
Db	901	GATGGTTCGTTACCTTGGCGACCCCACTGTCTCAAGAAATTTTGTCTAAGGGACATGTG	960
Qy	961	CGCGTTAGACCATATAAAGAAACCAATACAAAACCAAGCGAAATCTGTTGATGCGAA	1020
Db	961	CGCGTTAGACCATATAAAGAAACCAATACAAAACCAAGCGAAATCTGTTGATGCGAA	1020
Qy	1021	TATACGTACAGTTTACTCCCTTAAACCCCTGTGATGACGATTTTCAGACAGGCTCTCAAGAT	1080
Db	1021	TATACGTACAGTTTACTCCCTTAAACCCCTGTGATGACGATTTTCAGACAGGCTCTCAAGAT	1080
Qy	1081	ACTAAGCTATTGAAACACATAGCTATCGGTGACACCATCACTCTCAAGAAATTTACTAGCT	1140
Db	1081	ACTAAGCTATTGAAACACATAGCTATCGGTGACACCATCACTCTCAAGAAATTTACTAGCT	1140
Qy	1141	CAAGCAACAAAGCATTTTAAACAAAAACCCAGGCTTATACGATTTTATGAACGTGACTCC	1200
Db	1141	CAAGCAACAAAGCATTTTAAACAAAAACCCAGGCTTATACGATTTTATGAACGTGACTCC	1200
Qy	1201	TCAATCGTCACTCATGACCAATGACATTTTCCGTACGATTTTACCAATGAGTCAAGAGTTT	1260
Db	1201	TCAATCGTCACTCATGACCAATGACATTTTCCGTACGATTTTACCAATGAGTCAAGAGTTT	1260
Qy	1261	ACTTACCGTGTAAAAAATCGGGAAACAGCTTTATAGGATCAATAAAAAATCTGGTCTGAAT	1320
Db	1261	ACTTACCGTGTAAAAAATCGGGAAACAGCTTTATAGGATCAATAAAAAATCTGGTCTGAAT	1320
Qy	1321	GAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATACGTCCTTTAAAAAGGGGAA	1380
Db	1321	GAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATATACGTCCTTTAAAAAGGGGAA	1380
Qy	1381	AAGCCGTATGATCCCTTTTGTATCGCAGTCACTTCAAACTGTTCACCATCAATACGTTGAT	1440
Db	1381	AAGCCGTATGATCCCTTTTGTATCGCAGTCACTTCAAACTGTTCACCATCAATACGTTGAT	1440
Qy	1441	GTCGATACCAACGAATTTGCTTAAAAAGTGAGCAGCTCTTAA CAGCTAGCGAACGTAACCTTA	1500
Db	1441	GTCGATACCAACGAATTTGCTTAAAAAGTGAGCAGCTCTTAA CAGCTAGCGAACGTAACCTTA	1500
Qy	1501	GACTTCAGAGATTTATACGATCTCTGATGATAAGGCTAAACTACTCTTCAACAAATCTCGAT	1560
Db	1501	GACTTCAGAGATTTATACGATCTCTGATGATAAGGCTAAACTACTCTTCAACAAATCTCGAT	1560
Qy	1561	GCTTTTGGTATTATGACTATACCTTTAACTGGAAGAGTAGAGGATATACAGTACACACC	1620
Db	1561	GCTTTTGGTATTATGACTATACCTTTAACTGGAAGAGTAGAGGATATACAGTACACACC	1620
Qy	1621	AACCGTATCAATAACCGTTTATATGGCGAAGCGACCCGAGAGAGAGATGCTAGCTATCAT	1680
Db	1621	AACCGTATCAATAACCGTTTATATGGCGAAGCGACCCGAGAGAGAGATGCTAGCTATCAT	1680
Qy	1681	TTAGCTTATGATTAAGATCGTTTATACCGAAGAGACGAGAGGTTTACAGTACCTCGGT	1740
Db	1681	TTAGCTTATGATTAAGATCGTTTATACCGAAGAGACGAGAGGTTTACAGTACCTCGGT	1740
Qy	1741	TATACAGGAGACCTTATACCTGTATTAACCTTAAACGACAAATAA	1782
Db	1741	TATACAGGAGACCTTATACCTGTATTAACCTTAAACGACAAATAA	1782

RESULT 2
AAA37643
ID AAA37643 standard; DNA; 2096 BP.
XX
AC AAA37643;
XX
XX 15-SEP-2003 (revised)
DT 13-OCT-2000 (first entry)
XX
XX Chimeric SK-FBD coding sequence.
XX
XX Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
KW plasminogen; human; fibrinectin; thrombolytic therapy;
KW cardiovascular disorder; ss.
XX
XX Streptococcus dysgalactiae subsp. equisimilis.
OS Homo sapiens.
OS Chimeric.
XX
XX EP1024192-A2.
XX
XX 02-AUG-2000.
XX
XX 23-DEC-1999; 99EP-00310541.
XX
XX 24-DEC-1998; 98IN-DE003825.
XX
XX (COUL) CSIR COUNCIL SCI IND RES.
XX
XX Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
PI Yadav M;
PI
XX WPI; 2000-516032/47.
XX
XX Hybrid streptokinase-fibrin binding domain polypeptides useful for
PT thrombolytic therapy comprises a streptokinase fused with fibrin binding
PT domains of human fibrinectin.
XX
XX Example 6; Fig 22b; 58pp; English.
XX
XX This sequence represents a chimeric streptokinase-fibrin binding domain
CC (SK-FBD) protein coding sequence. The invention relates to a hybrid
CC plasminogen activator (PA) comprises a polypeptide fusion between
CC streptokinase (SK), which are capable of plasminogen (PG) activation, and
CC fibrin binding regions of human fibrinectin, which are from fibrin
CC binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA possesses the
CC ability to bind with fibrin independently and also characteristically
CC retains a PG activation ability which becomes evident only after a
CC pronounced duration, or lag, after exposure of the PA to a suitable
CC animal or human PG. The hybrid streptokinase-fibrin binding domain
CC polypeptides are useful in thrombolytic therapy for various kinds of
CC cardiovascular disorders. The hybrids have enhanced fibrin selectivity as
CC well as kinetics of plasminogen activation that are distinct from that of
CC natural streptokinase in being characterised by a temporary delay, or lag
CC of several minutes in the natural rate of the catalytic conversion of
CC plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins
CC can bind tightly with fibrin in blood clots soon after introduction into
CC the vascular system without significantly activating the circulating
CC blood plasminogen to plasmin, thus aiding in the localisation of the
CC plasminogen activation process to the site of pathological thrombus. This
CC overcomes systemic plasminogen activation encountered during clinical use
CC of streptokinase. (Updated on 15-SEP-2003 to standardise OS field)
XX
XX Sequence 2096 BP; 643 A; 480 C; 484 G; 489 T; 0 U; 0 Other;
SQ
Query Match 94.5%; Score 1684; DB 3; Length 2096;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1687; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
XX
XX 1 TCGCTTACGTTTCGGTATCGGTGATTCTTCTGTAACCAAGTAAAGCAACCCCGC 60
|||||

Db 51 TCGCTTACGTTTCGGTATCGGTGATTCTTCTGTAACCAAGTAAAGCAACCCCGC 110
Qy 61 CAGCTAGCCGGTCTCTCAACGACAGGAGCAGATCATCGCGACCCGTCGCGAGCAACCA 120
Db 111 CAGCTAGCCGGTCTCTCAACGACAGGAGCAGATCATCGCGACCCGTCGCGAGCAACCA 170
Qy 121 ACGTGGCCGAGATCTCGATCCCGCGAAATTAATACGACTCTATAGGAGACACCAAC 180
Db 171 ACGTGGCCGAGATCTCGATCCCGCGAAATTAATACGACTCTATAGGAGACACCAAC 230
Qy 181 GGTTCCTCTAGAAATAATTTTGTAACTTTTAAGAGGAGATATACCATGGTGAAGC 240
Db 231 GGTTCCTCTAGAAATAATTTTGTAACTTTTAAGAGGAGATATACCATGGTGAAGC 290
Qy 241 ACAACAGATTGTACCATAGCTGAGAAGTGTTCATCATGCTGCTGGAGCTTCCTATGT 300
Db 291 ACAACAGATTGTACCATAGCTGAGAAGTGTTCATCATGCTGCTGGAGCTTCCTATGT 350
Qy 301 GGTGGAGAAACGTGGGAGAGGCGAGCGACATCACTTGCACTTTAGAAATAGATGC 360
Db 351 GGTGGAGAAACGTGGGAGAGGCGAGCGACATCACTTGCACTTTAGAAATAGATGC 410
Qy 361 AACGATCAGGACACAGGACATCTTATAGATTGAGACACCTGGAGCAAGAGATAT 420
Db 411 AACGATCAGGACACAGGACATCTTATAGATTGAGACACCTGGAGCAAGAGATAT 470
Qy 421 CGAGAAACCTGCTCCAGTGCATCTGCACAGCAACGCGAGAGAGTGGAGTGTGAG 480
Db 471 CGAGAAACCTGCTCCAGTGCATCTGCACAGCAACGCGAGAGAGTGGAGTGTGAG 530
Qy 481 AGGCACACCTCTGTGCAGACACATCTGCGCGGATCTGCGCCCTTCAACGATGTCGTATT 540
Db 531 AGGCACACCTCTGTGCAGACACATCTGCGCGGATCTGCGCCCTTCAACGATGTCGTATT 590
Qy 541 GCTGGACCTGAGTGGTCTAGACCGTCCATCTGTCAACACAGCAATGGTGTGAGC 600
Db 591 GCTGGACCTGAGTGGTCTAGACCGTCCATCTGTCAACACAGCAATGGTGTGAGC 650
Qy 601 GTTCTGTGTTCTGAGGGAGCAATCAAGACATTAGTCTTAATTTTTTGAATTCAT 660
Db 651 GTTCTGTGTTCTGAGGGAGCAATCAAGACATTAGTCTTAATTTTTTGAATTCAT 710
Qy 661 CTAACATCAGCACCTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAGTCAAAAATCAAAA 720
Db 711 CTAACATCAGCACCTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAGTCAAAAATCAAAA 770
Qy 721 CCAATTTGCTACTGATAGTGGCGCATGTGCATAACTTGAGAAAGTGAATTAAG 780
Db 771 CCAATTTGCTACTGATAGTGGCGCATGTGCATAACTTGAGAAAGTGAATTAAG 830
Qy 781 GCTATTCAGAAACAAATTGATCGCTTAAACGTCACAGTAAAGAGACTACTTTGAGTCAAT 840
Db 831 GCTATTCAGAAACAAATTGATCGCTTAAACGTCACAGTAAAGAGACTACTTTGAGTCAAT 890
Qy 841 GATTTTCAAGCGATGCAACCATTAATCTGATCGAAACGCAAGGCTTACTTTGCTGACAAA 900
Db 891 GATTTTCAAGCGATGCAACCATTAATCTGATCGAAACGCAAGGCTTACTTTGCTGACAAA 950
Qy 901 GATGGTTCGGTAAACCTTGGCCGACCAACCTGTCGAAGAAATTTTGTGAAGCGGACATGTG 960
Db 951 GATGGTTCGGTAAACCTTGGCCGACCAACCTGTCGAAGAAATTTTGTGAAGCGGACATGTG 1010
Qy 961 CGGTTTGAACCATATAAAGAAAAACCAATACAAAAACCAAGCGAAATCTGTTGATGTGAA 1020
Db 1011 CGGTTTGAACCATATAAAGAAAAACCAATACAAAAACCAAGCGAAATCTGTTGATGTGAA 1070
Qy 1021 TATAGTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTCAAGAGCTCTCAAGAT 1080
Db 1071 TATAGTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTCAAGAGCTCTCAAGAT 1130
Qy 1081 ACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCAATCTCAAGAAATTAAGTACT 1140
Db 1131 ACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCAATCTCAAGAAATTAAGTACT 1190

QY 1141 CAAGCACAAGCATTTTAAACAAAACACCCAGGCTATACGATTTATGAACGCTACTCC 1200
 DB |||||
 1191 CAAGCACAAGCATTTTAAACAAAACACCCAGGCTATACGATTTATGAACGCTACTCC 1250
 QY 1201 TCAATCGTCATCATGACATGACATTTTCCGTAACGATTTTACCAATGGAATCAAGATTT 1260
 DB |||||
 1251 TCAATCGTCATCATGACATGACATTTTCCGTAACGATTTTACCAATGGAATCAAGATTT 1310
 QY 1261 ACTTACCGTGTAAATAACGGAACAGCTTATAGGATCAATAAATAATCTGCTGAAT 1320
 DB |||||
 1311 ACTTACCGTGTAAATAACGGAACAGCTTATAGGATCAATAAATAATCTGCTGAAT 1370
 QY 1321 GAAGAAATAAACAACACACTGACCTGATCTCTGAGAAATATTACGCTTAAATAAGGGAA 1380
 DB |||||
 1371 GAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGCTTAAATAAGGGAA 1430
 QY 1381 AAGCGTATGATCCCTTTGATCGAGTCACTTTGAAACTGTTCAACCATCAATACGTTGAT 1440
 DB |||||
 1431 AAGCGTATGATCCCTTTGATCGAGTCACTTTGAAACTGTTCAACCATCAATACGTTGAT 1490
 QY 1441 GTCGATACCAAGCAATGCTTAAAGTAGAGCTCTTAAACGCTAGCGAAACGTAACCTTA 1500
 DB |||||
 1491 GTCGATACCAAGCAATGCTTAAAGTAGAGCTCTTAAACGCTAGCGAAACGTAACCTTA 1550
 QY 1501 GACTTCAGAGATTTATAGATCTCTGATAGGCTAAACTACTCTACACATCTCGAT 1560
 DB |||||
 1551 GACTTCAGAGATTTATAGATCTCTGATAGGCTAAACTACTCTACACATCTCGAT 1610
 QY 1561 GCTTTGGTATTATGAGCTATACCTTAACTGGAAGTAGAGGATTAATCAGCATGACACC 1620
 DB |||||
 1611 GCTTTGGTATTATGAGCTATACCTTAACTGGAAGTAGAGGATTAATCAGCATGACACC 1670
 QY 1621 AACCGTATCAACCGTTTATATGGCAAGCGACCGAAGCGAGAGAGATGCTAGCTATCAT 1680
 DB |||||
 1671 AACCGTATCAACCGTTTATATGGCAAGCGACCGAAGCGAGAGAGATGCTAGCTACCAT 1730
 QY 1681 TTAGCTATGAT 1692
 DB |||||
 1731 TTAGCTGGTGT 1742

RESULT 3

AX80497
 ID AX80497 standard; cDNA; 2385 BP.
 XX AC
 AC AX80497;
 XX DT
 DT 26-AUG-1999 (first entry)
 XX DE
 DE Streptokinase and maltose binding protein fusion protein encoding cDNA.
 XX KW
 KW Streptococcus; streptokinase; fibrin-dependent plasminogen activator;
 KW nSK; rSK; bacterial; blood clot; thrombotic condition;
 KW myocardial infarction; venous thrombosis; pulmonary embolism;
 KW cerebral thrombosis; graft thrombosis; arterial thrombosis; ss.
 XX OS
 OS Streptococcus dysgalactiae subsp. equisimilis.
 XX OS
 OS Synthetic.
 XX PN
 PN W09931247-A1.
 XX XX
 XX 24-JUN-1999.
 XX XX
 XX 15-DEC-1998; 98WO-US026694.
 XX XX
 XX 15-DEC-1997; 97US-0069497P.
 XX XX
 XX (HARD) HARVARD COLLEGE.
 XX XX
 XX PI
 PI Reed GL;
 XX XX
 XX WPI; 1999-395183/33.

DR P-PSDB; AAY24797.
 XX N-terminally deleted streptokinase.
 PT
 XX
 PS Example; Page 45-48; 73pp; English.
 XX
 CC The present invention describes an isolated bacterial protein that
 CC induces fibrin-dependent plasminogen activation in a pharmaceutical
 CC composition for dissolving blood clots. Also described are: (1) a
 CC composition comprising an isolated modified streptokinase, the
 CC modification being removal of amino acid residues in the amino terminus;
 CC (2) a method for dissolving a blood clot in a subject, comprising
 CC administering to the subject a fibrin-dependent streptokinase protein; a
 CC nucleic acid (1) encoding a modified bacterial streptokinase; (3) an
 CC expression vector comprising (1); and (4) a host cell transformed with
 CC the expression vector of (3). The pharmaceutical composition comprising a
 CC bacterial fibrin-dependent plasminogen activator is useful for dissolving
 CC blood clots in patients with a thrombotic condition, e.g. myocardial
 CC infarction, venous thrombosis, pulmonary embolism, cerebral thrombosis,
 CC graft thrombosis and arterial thrombosis. The modified streptokinase can
 CC also be used in non-human mammals. Streptokinase activation of
 CC plasminogen is at least 10-fold, preferably 100-fold greater in the
 CC presence of fibrin than in the absence of fibrin. The modified
 CC streptokinase has at least one amino acid substitution that inactivates a
 CC substrate site for proteolytic cleavage. This reduces the rate of
 CC degradation of the streptokinase at least two-fold. The present sequence
 CC encodes a streptokinase and maltose binding protein fusion protein from
 CC an example of the present invention
 XX
 SQ Sequence 2385 BP; 745 A; 547 C; 549 G; 544 T; 0 U; 0 Other;
 Query Match 69.9%; Score 1245.8; DB 2; Length 2385;
 Best Local Similarity 99.4%; Pred. No. 0;
 Matches 1250; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 523 TTCAACCGATGTTTGGTATGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAC 582
 DB |||||
 1129 TCATCGAGGGTAGGATGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAC 1188
 QY 583 AGCCAAATGGTTTGTAGCGTTGCTGTTGAGGGGAGCAATCAAGACATTAAGTCTT 642
 DB |||||
 1189 AGCCAAATGTTTGTAGCGTTGCTGTTGAGGGGAGCAATCAAGACATTAAGTCTT 1248
 QY 643 AAATTTTGAATTCGATCTAAACATCAGACCTGCTCATCGAGGAAGACAGACGAGGC 702
 DB |||||
 1249 AAATTTTGAATTCGATCTAAACATCAGACCTGCTCATCGAGGAAGACAGACGAGGC 1308
 QY 703 TTAAGTCCAAATCAAAACCCATTTGCTACTGATAGTGGCGCGATGTCCACATAAACTTGAG 762
 DB |||||
 1309 TTAAGTCCAAATCAAAACCCATTTGCTACTGATAGTGGCGCGATGTCCACATAAACTTGAG 1368
 QY 763 AAAGCTGACTTACTAAAGGCTATTCAAGAAACAATTGATCGCTAACGTCACAGTAACGAC 822
 DB |||||
 1369 AAAGCTGACTTACTAAAGGCTATTCAAGAAACAATTGATCGCTAACGTCACAGTAACGAC 1428
 QY 823 GACTACTTTGAGGTCAATTTGATTTTGCAGCGATGCAACCAATTACTGTGCAAAACGGAAG 882
 DB |||||
 1429 GACTACTTTGAGGTCAATTTGATTTTGCAGCGATGCAACCAATTACTGTGCAAAACGGAAG 1488
 QY 883 GTCTACTTTGCTGACAAAGATGTTGCTGTAACCTTGCAGCCCAACCTGTCGAAGAATTT 942
 DB |||||
 1489 GTCTACTTTGCTGACAAAGATGTTGCTGTAACCTTGCAGCCCAACCTGTCGAAGAATTT 1548
 QY 943 TTGCTAAGCGGACATGTCGCGGTTAGACCATATAAAGAAAAAACAATACAAAACCAAGCG 1002
 DB |||||
 1549 TTGCTAAGCGGACATGTCGCGGTTAGACCATATAAAGAAAAAACAATACAAAACCAAGCG 1608
 QY 1003 AAATCTGTTGATGTGGAATATATCTGTACAGTTTACTCCCTTAAACCTGTATGAGATTTC 1062
 DB |||||
 1609 AAATCTGTTGATGTGGAATATATCTGTACAGTTTACTCCCTTAAACCTGTATGAGATTTC 1668
 QY 1063 AGACCAAGGTCTCAAGATCTAAGCTATTGAAACACTAGCTATCGGTGACCATCACA 1122
 DB |||||

Db 1669 AGACAGGTCTCAAGATACCTAAGCTATTGAACACTAGCTATCGGTGACACCATCACA 1728
Qy 1123 TCTCAAGAAATTACTAGCTCAAGCACAAGCAATTTTAAACAAAACCAACCCAGGCTATACG 1182
Db 1729 TCTCAAGAAATTACTAGCTCAAGCACAAGCAATTTTAAACAAAACCAACCCAGGCTATACG 1788
Qy 1183 ATTTATGAACGTGACTCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTA 1242
Db 1789 ATTTATGAACGTGACTCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTA 1848
Qy 1243 CCAATGGATCAAGAGTTTACTTACCGTGTATAAATCGGGAACAAGCTTATAGATCAAT 1302
Db 1849 CCATATGGATCAAGAGTTTACTTACCGTGTATAAATCGGGAACAAGCTTATAGATCAAT 1908
Qy 1303 AAAAAATCTGGTCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAC 1362
Db 1909 AAAAAATCTGGTCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAC 1968
Qy 1363 GTCCTTAAAAAGGGGAAAGCGGTATGATCCCTTTGATGCGAGTCACTTTGAAACTGTTTC 1422
Db 1969 GTCCTTAAAAAGGGGAAAGCGGTATGATCCCTTTGATGCGAGTCACTTTGAAACTGTTTC 2028
Qy 1423 ACCATCAATAGCTTGATCTGATACCAACGAAATTCGTAAGAGTGAAGCTCTTAACA 1482
Db 2029 ACCATCAATAGCTTGATCTGATACCAACGAAATTCGTAAGAGTGAAGCTCTTAACA 2088
Qy 1483 GCTAGCGAACGTAACCTTAGACTTCAGAGATTTATACGATCCTCGTGAAGGCTTAAACTA 1542
Db 2089 GCTAGCGAACGTAACCTTAGACTTCAGAGATTTATACGATCCTCGTGAAGGCTTAAACTA 2148
Qy 1543 CTCTACAAACATCTCGATGCTTTTGGTATTATGGACTATACCTTAACCTGGAAGTAGAG 1602
Db 2149 CTCTACAAACATCTCGATGCTTTTGGTATTATGGACTATACCTTAACCTGGAAGTAGAG 2208
Qy 1603 GATAATCAGCATGACACCAACCGTATCATAAACCGTTTATATGGGCAAGCACCAGGA 1662
Db 2209 GATAATCAGCATGACACCAACCGTATCATAAACCGTTTATATGGGCAAGCACCAGGA 2268
Qy 1663 GAGATGCTAGCTATCATTTAGCTATGATAAGATCGTTATACCGAAGAACGAGAA 1722
Db 2269 GAGATGCTAGCTATCATTTAGCTATGATAAGATCGTTATACCGAAGAACGAGAA 2328
Qy 1723 GTTTACAGTACTCGTGGTTATACAGGACACCTTATACCTGATTAACCTTACGACAAA 1779
Db 2329 GTTTACAGTACTCGTGGTTATACAGGACACCTTATACCTGATTAACCTTACGACAAA 2385

RESULT 4
AAA37622
ID AAA37622 standard; DNA; 1377 BP.
XX AC AAA37622;
XX 15-SEP-2003 (revised)
DT 13-OCT-2000 (first entry)
XX DE Streptokinase-NTRN gene.
XX Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
KW plasminogen; human; fibrinectin; thrombolytic therapy;
KW cardiovascular disorder; ss.
XX Streptococcus dysgalactiae subsp. equisimilis.
OS EP1024192-A2.
XX FN 02-AUG-2000.
XX PD 23-DEC-1999; 99EP-00310541.
XX PF 24-DEC-1998; 98IN-DE003825.
XX PR (COUL) CSIR COUNCIL SCI IND RES.

XX Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
PI Yadav M;
XX WPI; 2000-516032/47.
XX Hybrid streptokinase-fibrin binding domain polypeptides useful for
PT thrombolytic therapy comprises a streptokinase fused with fibrin binding
PT domains of human fibrinectin.
XX Example 1; Fig 11; 58pp; English.
PS This sequence represents a streptokinase-NTRN (SK-NTRN) gene (where NTRN
XX stands for N-terminally repaired with native sequence). The invention
CC relates to a hybrid plasminogen activator (PA) comprises a polypeptide
CC fusion between streptokinase (SK), which are capable of plasminogen (PG)
CC activation, and fibrin binding regions of human fibrinectin, which are
CC from fibrin binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA
CC possesses the ability to bind with fibrin independently and also
CC characteristically retains a PG activation ability which becomes evident
CC only after a pronounced duration, or lag, after exposure of the PA to a
CC suitable animal or human PG. The hybrid streptokinase-fibrin binding
CC domain polypeptides are useful in thrombolytic therapy for various kinds
CC of cardiovascular disorders. The hybrids have enhanced fibrin selectivity
CC as well as kinetics of plasminogen activation that are distinct from that
CC of natural streptokinase in being characterised by a temporary delay, or
CC lag of several minutes in the natural rate of the catalytic conversion of
CC plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins
CC can bind tightly with fibrin in blood clots soon after introduction into
CC the vascular system without significantly activating the circulating
CC blood plasminogen to plasmin, thus aiding in the localisation of the
CC plasminogen activation process to the site of pathological thrombus. This
CC overcomes systemic plasminogen activation encountered during clinical use
CC of streptokinase. (Updated on 15-SEP-2003 to standardise OS field)
XX SQ Sequence 1377 BP; 465 A; 302 C; 265 G; 345 T; 0 U; 0 Other;

Query Match 69.9%; Score 1245; DB 3; Length 1377;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 538 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCTGTCATCTGTCAACACAGCAATGCTGTT 597
Db 133 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCTGTCATCTGTCAACACAGCAATGCTGTT 192
Qy 598 AGCGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 657
Db 193 AGCGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 252
Qy 658 GATCTAATCATCAGACCTGCTCATGGAGGAAGACAGACGAGGCTTAAGTCCAAATCA 717
Db 253 GATCTAATCATCAGACCTGCTCATGGAGGAAGACAGACGAGGCTTAAGTCCAAATCA 312
Qy 718 AAACCATTTGCTACTGATAGTGGCGCATGTGCATATAAATTTGAGAAGCTGACTTACTA 777
Db 313 AAACCATTTGCTACTGATAGTGGCGCATGTGCATATAAATTTGAGAAGCTGACTTACTA 372
Qy 778 AAGGCTATTCAAGAACAAATGATGCTTAACGTTCCACAGTAAACGACGACTACTTTGAGTTC 837
Db 373 AAGGCTATTCAAGAACAAATGATGCTTAACGTTCCACAGTAAACGACGACTACTTTGAGTTC 432
Qy 838 ATTGATTTTGCAGCGATGCAACCATTTACTGATCGAAACGCAAGGCTTACTTTGCTGAC 897
Db 433 ATTGATTTTGCAGCGATGCAACCATTTACTGATCGAAACGCAAGGCTTACTTTGCTGAC 492
Qy 898 AAGAGTGGTTCCGTTAACTTTGCCGACCCCAACCTGTCACAGAAATTTTTCGTAAGCGGACAT 957
Db 493 AAGAGTGGTTCCGTTAACTTTGCCGACCCCAACCTGTCACAGAAATTTTTCGTAAGCGGACAT 552
Qy 958 GTGCGGCTTAGACCATATAAAGAAAAACCAATACAAAAACCAAGGAAATCTGTTGATGTG 1017
Db 553 GTGCGGCTTAGACCATATAAAGAAAAACCAATACAAAAACCAAGGAAATCTGTTGATGTG 612

QY 1018 GAATATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGTCTCAAA 1077
 Db 613 GAATATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGTCTCAAA 672
 QY 1078 GATACCTAGCTATTGAACACCTAGCTATCGGTGACACCATCATCTCAGAAATTACTA 1137
 Db 673 GATACCTAGCTATTGAACACCTAGCTATCGGTGACACCATCATCTCAGAAATTACTA 732
 QY 1138 GCTCAAGCACAAGCAATTTTAAACAAAACCCAGGCTATACGATTTTATGAACGTGAC 1197
 Db 733 GCTCAAGCACAAGCAATTTTAAACAAAACCCAGGCTATACGATTTTATGAACGTGAC 792
 QY 1198 TCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 1257
 Db 793 TCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 852
 QY 1258 TTTACTTACCGTGTAAATATCGGGAACAGCTTATAGATCAATAAAAATCTGCTCTG 1317
 Db 853 TTTACTTACCGTGTAAATATCGGGAACAGCTTATAGATCAATAAAAATCTGCTCTG 912
 QY 1318 AATGAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTTACGCTCTTAAAAAGGG 1377
 Db 913 AATGAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTTACGCTCTTAAAAAGGG 972
 QY 1378 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCACCATCAATATACGTT 1437
 Db 973 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCACCATCAATATACGTT 1032
 QY 1438 GATGTCGATACCAAGAAATGCTTAAAGTAGAGCAGCTTTTAAACAGCTAGCGGAACCTAAC 1497
 Db 1033 GATGTCGATACCAAGAAATGCTTAAAGTAGAGCAGCTTTTAAACAGCTAGCGGAACCTAAC 1092
 QY 1498 TTAGACTTCAGAGATTTATACGATCTCTGATAGGCTTAAAGCTTAACTCTACACATCTC 1557
 Db 1093 TTAGACTTCAGAGATTTATACGATCTCTGATAGGCTTAAAGCTTAACTCTACACATCTC 1152
 QY 1558 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAGAGTAGAGGATATACAGATGAC 1617
 Db 1153 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAGAGTAGAGGATATACAGATGAC 1212
 QY 1618 ACCAACCGTATCAATACCGTTTATATGGCAAGCGACCCGAGAGAGAGATGCTAGCTAT 1677
 Db 1213 ACCAACCGTATCAATACCGTTTATATGGCAAGCGACCCGAGAGAGAGATGCTAGCTAT 1272
 QY 1678 CATTTAGCTTATGATAAGATCGTTATACGGAAGAGACGAGAGTTTACAGTACCTG 1737
 Db 1273 CATTTAGCTTATGATAAGATCGTTATACGGAAGAGACGAGAGTTTACAGTACCTG 1332
 QY 1738 CGTTATACAGGACACCTTACTGATACCCCTTAACGACAAATAA 1782
 Db 1333 CGTTATACAGGACACCTTACTGATACCCCTTAACGACAAATAA 1377

RESULT 5
 AAA37633 standard; DNA; 1245 BP.

AC AAA37633;
 XX
 XX
 DT 15-SEP-2003 (revised)
 DT 13-OCT-2000 (first entry)

DE S. equisimilis streptokinase coding sequence.

XX Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
 KW plasminogen; human; fibrinectin; thrombolytic therapy;
 KW cardiovascular disorder; fibrinectin; ss.

XX Streptococcus dysgalactiae subsp. equisimilis.

XX Key Location/Qualifiers
 FH 1..1245
 FT CDS /*tag= a

FT
 XX
 PN
 XX
 XX
 PD
 XX
 XX
 PF
 XX
 XX
 PR
 XX
 XX
 PA
 XX
 PI
 PI
 Y
 Y
 DR
 DR
 XX
 XX
 PT
 PT
 XX
 PS
 XX

/product= "streptokinase"

EP1024192-A2.
 02-AUG-2000.
 23-DEC-1999; 99EP-00310541.
 24-DEC-1998; 98IN-DE003825.
 (COUL) CSIR COUNCIL SCI IND RES.

Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
 Yadav M;
 WPI; 2000-516032/47.
 P-PSDB; AAY90282.

Hybrid streptokinase-fibrin binding domain polypeptides useful for
 thrombolytic therapy comprises a streptokinase fused with fibrin binding
 domains of human fibrinectin.

Example 3; Fig 3; 58pp; English.

This sequence represents the human Streptococcus equisimilis
 streptokinase coding sequence. The invention relates to a hybrid
 plasminogen activator (PA) comprises a polypeptide fusion between
 streptokinase (SK), which are capable of plasminogen (PG) activation, and
 fibrin binding regions of human fibrinectin, which are from fibrin
 binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA possesses the
 ability to bind with fibrin independently and also characteristically
 retains a PG activation ability which becomes evident only after a
 pronounced duration, or lag, after exposure of the PA to a suitable
 animal or human PG. The hybrid streptokinase-fibrin binding domain
 polypeptides are useful in thrombolytic therapy for various kinds of
 cardiovascular disorders. The hybrids have enhanced fibrin selectivity as
 well as kinetics of plasminogen activation that are distinct from that of
 natural streptokinase in being characterised by a temporary delay, or lag
 of several minutes in the natural rate of the catalytic conversion of
 plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins
 can bind tightly with fibrin in blood clots soon after introduction into
 the vascular system without significantly activating the circulating
 blood plasminogen to plasmin, thus aiding in the localisation of the
 plasminogen activation process to the site of pathological thrombus. This
 overcomes systemic plasminogen activation encountered during clinical use
 of streptokinase. (Updated on 15-SEP-2003 to standardise OS field)

Sequence 1245 BP; 426 A; 267 C; 237 G; 315 T; 0 U; 0 Other;
 Query Match 69.8%; Score 1243.4; DB 3; Length 1245;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1244; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 538 ATTGCTGGACCTCAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAAATGGTTGTT 597
 Db 1 ATTGCTGGACCTCAGTGGCTGTAGACCGTCCATCTGTCTCAACACAGCCAAATGGTTGTT 60
 QY 598 AGCGTTGCTGGTACTGTGTGAGGGGACGAATCAAGACATTTAGTCTTAAATTTTGAATC 657
 Db 61 AGCGTTGCTGGTACTGTGTGAGGGGACGAATCAAGACATTTAGTCTTAAATTTTGAATC 120
 QY 658 GATCTAACATCAGACCTGCTCATGGAGGAAAGACAGAGAGCTTTAAGTCCAAATCA 717
 Db 121 GATCTAACATCAGACCTGCTCATGGAGGAAAGACAGAGAGCTTTAAGTCCAAATCA 180
 QY 718 AAACCAATTTGCTACTCATAGTGGCGGATGTACATAAACTTGAGAAGCTGACTTACTA 777
 Db 181 AAACCAATTTGCTACTCATAGTGGCGGATGTACATAAACTTGAGAAGCTGACTTACTA 240
 QY 778 AAGGCTATTCAAGACAATTTGATCGCTTAACGTCACAGTAAACGACGACTACTTTGAGGTC 837
 Db 241 AAGGCTATTCAAGACAATTTGATCGCTTAACGTCACAGTAAACGACGACTACTTTGAGGTC 300

QY 838 ATTGATTTTGCAGCGATGCAACCAATTAATGATCGAAACGCAAGGTCTACTTTGCTGAC 897
DB |||||
301 ATTGATTTTGCAGCGATGCAACCAATTAATGATCGAAACGCAAGGTCTACTTTGCTGAC 360
QY 898 AAGATGGTTTGGTAACTTGGCGACCAACCTGTCCAGAAATTTTCTAAGCGGACAT 957
DB |||||
361 AAGATGGTTTGGTAACTTGGCGACCAACCTGTCCAGAAATTTTCTAAGCGGACAT 420
QY 958 GTGCGGTTTAGACCATATAAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 1017
DB |||||
421 GTGCGGTTTAGACCATATAAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 480
QY 1018 GAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTACAGACGCTCTCAAA 1077
DB |||||
481 GAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTACAGACGCTCTCAAA 540
QY 1078 GATACCTAAGCTATTGAACACCTAGCTATCGGTGACCAATCATCATCTCAAGAAATTAATA 1137
DB |||||
541 GATACCTAAGCTATTGAACACCTAGCTATCGGTGACCAATCATCATCTCAAGAAATTAATA 600
QY 1138 GCTCAAGCAAAAGCAATTTTAAACAAAAACCCAGGCTATACGATTTTATGAACGTGAC 1197
DB |||||
601 GCTCAAGCAAAAGCAATTTTAAACAAAAACCCAGGCTATACGATTTTATGAACGTGAC 660
QY 1198 TCTCAATCGTCACTCATGACCAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 1257
DB |||||
661 TCTCAATCGTCACTCATGACCAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 720
QY 1258 TTTACTTACCGTGTAAATAATCGGAAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 1317
DB |||||
721 TTTACTTACCGTGTAAATAATCGGAAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 780
QY 1318 AATGAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1377
DB |||||
781 AATGAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 840
QY 1378 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTTACCAATCAATACGTT 1437
DB |||||
841 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTTACCAATCAATACGTT 900
QY 1438 GATGCGATACCAACGAATTTGCTTAAAGTGAGCAGCTCTTAAACAGCTACGCAACGTAAC 1497
DB |||||
901 GATGCGATACCAACGAATTTGCTTAAAGTGAGCAGCTCTTAAACAGCTACGCAACGTAAC 960
QY 1498 TTAGACTTCAGAGATTTATACGATCTCTGATGATAAGCTTAACTCTTCAACAATCTC 1557
DB |||||
961 TTAGACTTCAGAGATTTATACGATCTCTGATGATAAGCTTAACTCTTCAACAATCTC 1020
QY 1558 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAAGCTAGAGGATATACGATGAC 1617
DB |||||
1021 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAAGCTAGAGGATATACGATGAC 1080
QY 1618 ACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAGGAGAGATGCTAGCTAT 1677
DB |||||
1081 ACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAGGAGAGATGCTAGCTAT 1140
QY 1678 CATTTAGCTTATGATAAGATCGTTTATACCGAAGAGAACGAGAAGTTTACAGCTACCTG 1737
DB |||||
1141 CATTTAGCTTATGATAAGATCGTTTATACCGAAGAGAACGAGAAGTTTACAGCTACCTG 1200
QY 1738 CGTTATACGGACACCTATACCTGATACCCCTTAAACGACAAATAA 1782
DB |||||
1201 CGTTATACGGACACCTATACCTGATACCCCTTAAACGACAAATAA 1245

RESULT 6
AAQ11651
ID AAQ11651 standard; DNA; 2030 BP.
XX
AC AAQ11651;
XX
DT 08-JUL-1991 (first entry)
XX

DE FB-FB-SK fusion construct.
XX Fibrin-binding protein; fibrinolysis; intravascular thrombi; fibrinogen;
KW streptokinase; fusion protein; ss.
XX Staphylococcus aureus.
OS
XX
FH Key Location/Qualifiers
FT CDS 10..184
FT /*tag= a
FT /label= FB monomer
FT CDS 185..358
FT /*tag= b
FT /label= FB monomer
FT CDS 359..1601
FT /*tag= c
FT /label= streptokinase
XX
FN US5011686-A.
XX
PD 30-APR-1991.
XX
PF 15-NOV-1989; 89US-00437769.
XX
PR 21-SEP-1987; 87US-00099242.
XX
PA (CREA-) CREATIVE BIOMOLEC.
XX
PI Pang RHL;
XX
DR WPI; 1991-140198/19.
XX
DR P-PSDB; AAR11829.
XX
XX Imparting injectable fibrinolytic agent - with affinity for intravascular
PT thrombus, by linking agent to fibrin binding domain.
XX
PS Disclosure; Fig 5; 18pp; English.
XX
CC The DNA encodes an FB-PB dimer linked to the streptokinase coding
CC sequence. The FB fragment has selective affinity for fibrin, low affinity
CC for fibrinogen, and minimal immunogenicity, imparting thrombus-targeting
CC capability. See also AAQ11649 and AAQ11650
XX
SQ Sequence 2030 BP; 673 A; 464 C; 406 G; 487 T; 0 U; 0 Other;

Query Match 69.7%; Score 1241.8; DB 2; Length 2030;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1243; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 538 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCCAAATGGTGT 597
DB |||||
358 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCCAAATGGTGT 417
QY 598 AGCGTTGCTGCTACTGTTGAGGGGACCAATCAAGACATTAAGTCTTAAATTTTGAATC 657
DB |||||
418 AGCGTTGCTGCTACTGTTGAGGGGACCAATCAAGACATTAAGTCTTAAATTTTGAATC 477
QY 658 GATCTAACATCAACGACCTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAAGTCCAAATCA 717
DB |||||
478 GATCTAACATCAACGACCTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAAGTCCAAATCA 537
QY 718 AAACCAATTTGCTACTGATAGTGGCGGATGTACATAAACTTGAAGAGCTGACTTACTA 777
DB |||||
538 AAACCAATTTGCTACTGATAGTGGCGGATGTACATAAACTTGAAGAGCTGACTTACTA 597
QY 778 AAGGCTATTCAAGAAACAAATTTGATCGCTAAAGTCCACAGTAAACGACGACTTTGAGTGC 837
DB |||||
598 AAGGCTATTCAAGAAACAAATTTGATCGCTAAAGTCCACAGTAAACGACGACTTTGAGTGC 657
QY 838 ATTGATTTTGCAGCGATGCAACCAATTAATGATCGAAACGCAAGGTCTACTTTGCTGAC 897
DB |||||
658 ATTGATTTTGCAGCGATGCAACCAATTAATGATCGAAACGCAAGGTCTACTTTGCTGAC 717

QY 898 AAAGATGGTTCGGTAAACCTTGGCGACCCCAACCTGTCACAGAAATTTTGTCTAAAGCGGACAT 957
 Db 718 AAAGATGGTTCGGTAAACCTTGGCGACCCCAACCTGTCACAGAAATTTTGTCTAAAGCGGACAT 777
 QY 958 GTGCGGTTAGACCATATATAAGAAAAACCAATACAAAAACCAAGCGAAATCTGTTGATGTG 1017
 Db 778 GTGCGGTTAGACCATATATAAGAAAAACCAATACAAAAACCAAGCGAAATCTGTTGATGTG 837
 QY 1018 GAATATACGTGACGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAAA 1077
 Db 838 GAATATACGTGACGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAAA 897
 QY 1078 GATACCTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCTCAAGAAATTA 1137
 Db 898 GATACCTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCTCAAGAAATTA 957
 QY 1138 GCTCAAGCAAAAGCAATTTTAAACAAAAACCCAGGCTATACGATTTATGAAAGCTGAC 1197
 Db 958 GCTCAAGCAAAAGCAATTTTAAACAAAAACCCAGGCTATACGATTTATGAAAGCTGAC 1017
 QY 1198 TCCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAGATGATCAAGAG 1257
 Db 1018 TCCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAGATGATCAAGAG 1077
 QY 1258 TTTACTTACCGTGTGTTAAATAACCGGAAACAGCTTATAGGATCAATAAAAAATCTGCTCG 1317
 Db 1078 TTTACTTACCGTGTGTTAAATAACCGGAAACAGCTTATAGGATCAATAAAAAATCTGCTCG 1137
 QY 1318 AATGAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTAATGCTCTTAAAAAGGG 1377
 Db 1138 AATGAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTAATGCTCTTAAAAAGGG 1197
 QY 1378 GAAAGCCGTATGATCCCTTTCATGCGAGTCACTTGAACTGTTCCACATCAATAGCTT 1437
 Db 1198 GAAAGCCGTATGATCCCTTTCATGCGAGTCACTTGAACTGTTCCACATCAATAGCTT 1257
 QY 1438 GATGTCGATACCAACGAATGCTTAAAGTGAGCAGCTCTTAAAGCTAGCGAACGTAAC 1497
 Db 1258 GATGTCGATACCAACGAATGCTTAAAGTGAGCAGCTCTTAAAGCTAGCGAACGTAAC 1317
 QY 1498 TTAGACTTCAGAGATTTATACGATCTCTGATAGGCTTAACTTCTTCAACAATCTC 1557
 Db 1318 TTAGACTTCAGAGATTTATACGATCTCTGATAGGCTTAACTTCTTCAACAATCTC 1377
 QY 1558 GATGCTTTTGGTATTTATGACACTTACCTTAACTGGAAGAGTAGAGATATACGATGAC 1617
 Db 1378 GATGCTTTTGGTATTTATGACACTTACCTTAACTGGAAGAGTAGAGATATATCAGATGAC 1437
 QY 1618 ACCAACCGTATCATACCGTTTTATATGGCAAGCGACCCGAGGAGAGATGCTAGCTAT 1677
 Db 1438 ACCAACCGTATCATACCGTTTTATATGGCAAGCGACCCGAGGAGAGATGCTAGCTAT 1497
 QY 1678 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAGAGACGAGAAGTTTACAGCTACCTG 1737
 Db 1498 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAGAGACGAGAAGTTTACAGCTACCTG 1557
 QY 1738 CGTTATACGGGACACCTATACCTGATTAACCCCTTAACGACAAATPAA 1782
 Db 1558 CGTTATACGGGACACCTATACCTGATTAACCCCTTAACGACAAATPAA 1602

RESULT 7

AA80492
 ID AA80492 standard; cDNA; 1242 BP.

XX AC AA80492;

XX AC

DT 17-OCT-2003 (revised)

DT 26-AUG-1999 (first entry)

XX Streptococcus equisimilis native streptokinase encoding cDNA.

XX Streptococcus; streptokinase; fibrin-dependent plasminogen activator;

KW nSK; rSK; bacterial; blood clot; thrombotic condition;
 KW myocardial infarction; venous thrombosis; pulmonary embolism;
 KW cerebral thrombosis; graft thrombosis; arterial thrombosis; ss.
 XX Streptococcus dysgalactiae subsp. equisimilis.
 OS WO9931247-A1.
 XX 24-JUN-1999.
 XX 15-DEC-1998; 98WO-US026694.
 XX 15-DEC-1997; 97US-0069497P.
 XX (HARD) HARVARD COLLEGE.
 XX Reed GL;
 XX WPI; 1999-395183/33.
 DR P-PSDB; AAY24794.
 XX N-terminally deleted streptokinase.
 XX Claim 44; Page 58-60; 73pp; English.
 XX The present invention describes an isolated bacterial protein that induces fibrin-dependent plasminogen activation in a pharmaceutical composition for dissolving blood clots. Also described are: (1) a composition comprising an isolated modified streptokinase, the modification being removal of amino acid residues in the amino terminus; (2) a method for dissolving a blood clot in a subject, comprising administering to the subject a fibrin-dependent streptokinase protein; a nucleic acid (I) encoding a modified bacterial streptokinase; (3) an expression vector comprising (I); and (4) a host cell transformed with the expression vector of (3). The pharmaceutical composition comprising a bacterial fibrin-dependent plasminogen activator is useful for dissolving blood clots in patients with a thrombotic condition, e.g. myocardial infarction, venous thrombosis, pulmonary embolism, cerebral thrombosis, graft thrombosis and arterial thrombosis. The modified streptokinase can also be used in non-human mammals. Streptokinase activation of plasminogen is at least 10-fold, preferably 100-fold greater in the presence of fibrin than in the absence of fibrin. The modified streptokinase has at least one amino acid substitution that inactivates a substrate site for proteolytic cleavage. This reduces the rate of degradation of the streptokinase at least two-fold. The present sequence encodes native streptokinase (nSK). (Updated on 17-OCT-2003 to standardise OS field)

XX SQ Sequence 1242 BP; 424 A; 267 C; 237 G; 314 T; 0 U; 0 Other;

Query Match 69.6%; Score 1240.4; DB 2; Length 1242;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1241; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 538 ATTGCTGGACCTCAGTGGCTGTAGACCGTCCATCTCTCAACACAGCAATGCTGTT 597
 Db 1 ATTGCTGGACCTCAGTGGCTGTAGACCGTCCATCTCTCAACACAGCAATGCTGTT 60
 QY 598 AGCGTTGCTCGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTTAAATCTTGAATC 657
 Db 61 AGCGTTGCTCGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTTAAATCTTGAATC 120
 QY 658 GATCTAACATCATCGACCTGCTCATGGAGGAAAGACAGACAGAGCTTAAGTCCAAATCA 717
 Db 121 GATCTAACATCATCGACCTGCTCATGGAGGAAAGACAGACAGAGCTTAAGTCCAAATCA 180
 QY 718 AARACATTTGCTACTGATAGTGGCGGATGTCACATAAACTTGAGAGAGCTGACTACTA 777
 Db 181 AARACATTTGCTACTGATAGTGGCGGATGTCACATAAACTTGAGAGAGCTGACTACTA 240
 QY 778 AAGGCTATTCAAGAACAAATGATCGCTAAAGTCCACAGTAAACGACGACTACTTTGAGTGC 837
 Db 241 AAGGCTATTCAAGAACAAATGATCGCTAAAGTCCACAGTAAACGACGACTACTTTGAGTGC 300

```
QY 838 ATTGATTTTCCAAAGCGATGCAACCAATTAATGATCGAATCGGCAAGGCTCTACTTTGCTGAC 897
Db |||||
QY 301 ATTGATTTTCCAAAGCGATGCAACCAATTAATGATCGAATCGGCAAGGCTCTACTTTGCTGAC 360
Db |||||
QY 898 AAAGATGGTTTCGGTAAACCTTTCGGCCGACCAACCTCTCCAAAGAAATTTTTCGTAAGCGGACAT 957
Db |||||
QY 361 AAAGATGGTTTCGGTAAACCTTTCGGCCGACCAACCTCTCCAAAGAAATTTTTCGTAAGCGGACAT 420
QY 958 GTCCGGGTTAGACCATATAAGAAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 1017
Db |||||
QY 421 GTCCGGGTTAGACCATATAAGAAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 480
QY 1018 GAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGANTTTTCAGACCAAGGCTCTCAA 1077
Db |||||
QY 481 GAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGANTTTTCAGACCAAGGCTCTCAA 540
QY 1078 GATACCTAAGCTATTGAACCACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTA 1137
Db |||||
QY 541 GATACCTAAGCTATTGAACCACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTA 600
QY 1138 GCTCAAGCACAAGCAATTTTAAACAAAACCAACCCAGGCTATAGGATTTATGAACGTCAC 1197
Db |||||
QY 601 GCTCAAGCACAAGCAATTTTAAACAAAACCAACCCAGGCTATAGGATTTATGAACGTCAC 660
QY 1198 TCTCAATCGTCACTCATGCAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 1257
Db |||||
QY 661 TCTCAATCGTCACTCATGCAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 720
QY 1258 TTTACTTACCGTGTAAATAACCGGAAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 1317
Db |||||
QY 721 TTTACTTACCGTGTAAATAACCGGAAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 780
QY 1318 AATGAAGAATAAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGG 1377
Db |||||
QY 781 AATGAAGAATAAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGG 840
QY 1378 GAAACCGGTATGATCCCTTTGATCGCAGTCACTTTGAAACTGTTTCAACATCAATACGTT 1437
Db |||||
QY 841 GAAACCGGTATGATCCCTTTGATCGCAGTCACTTTGAAACTGTTTCAACATCAATACGTT 900
QY 1438 GATGTCGATACCAACGAATTTGCTTAAAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 1497
Db |||||
QY 901 GATGTCGATACCAACGAATTTGCTTAAAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 960
QY 1498 TTAGACTTCAGAGATTTATAGGATCTCTCGTGATAGGCTAAACTACTCTACAACTCTC 1557
Db |||||
QY 961 TTAGACTTCAGAGATTTATAGGATCTCTCGTGATAGGCTAAACTACTCTACAACTCTC 1020
QY 1558 GATGCTTTTGGTATTATGAGTACTTACCTTAACTGGAAGGTAGAGGATTAATCAGGATGAC 1617
Db |||||
QY 1021 GATGCTTTTGGTATTATGAGTACTTACCTTAACTGGAAGGTAGAGGATTAATCAGGATGAC 1080
QY 1618 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAGGAGAGATGCTAGCTAT 1677
Db |||||
QY 1081 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAGGAGAGATGCTAGCTAT 1140
QY 1678 CATTTAGCCTATGATAAAGATCGTTTATACCGAAGAGAAACGAGAAGTTTACAGTACCTG 1737
Db |||||
QY 1141 CATTTAGCCTATGATAAAGATCGTTTATACCGAAGAGAAACGAGAAGTTTACAGTACCTG 1200
QY 1738 CGTTATACAGGGACCTATACCTGATACCCCTTAACGACAAA 1779
Db |||||
QY 1201 CGTTATACAGGGACCTATACCTGATACCCCTTAACGACAAA 1242
```

RESULT 8

ABA05546
ID ABA05546 standard; cDNA; 1254 BP.

XX
AC ABA05546;
XX

DT 26-FEB-2002 (first entry)

Streptokinase cDNA.

Streptokinase; cerebroprotective; cardiact; gene therapy; fusion;
vasodilator; thrombolytic; angina; myocardial infarction; stroke;
gene therapy; maxadilan; ss.

Unidentified.

WO200185100-A2.

15-NOV-2001.

10-MAY-2001; 2001WO-US015209.

11-MAY-2000; 2000US-00569920.

(GEO) GEN HOSPITAL CORP.

Reddy VB, Lerner E;

WPI; 2002-062184/08.

New fusion protein or conjugate, useful for treating unstable angina,
acute myocardial infarction or stroke, comprises a vasodilator
polypeptide and a thrombolytic polypeptide, or active fragments of the
polypeptides.

Example 1; Fig 2; 37pp; English.

The invention relates to a fusion protein or a conjugate comprising a
vasodilator polypeptide, or its active fragment, and a thrombolytic
polypeptide or its active fragment. The protein is useful for treating a
subject suffering from a partially or totally occluded blood vessel,
causing unstable angina, acute myocardial infarction or stroke. The
polynucleotide encoding the polypeptide is useful in gene therapy. The
vasodilatory action of the protein allows for the use of lower doses of a
thrombolytic while maintaining the clot dissolving effectiveness of the
thrombolytic, and the use of lower doses of the thrombolytic reduces
associated side effects. The present sequence is the streptokinase cDNA
used in the construction of a Maxadilan-Streptokinase fusion protein.
Maxadilan is a vasodilator peptide produced by the salivary gland of the
New World sand fly

Sequence 1254 BP; 426 A; 271 C; 241 G; 316 T; 0 U; 0 Other;

Query Match 69.6%; Score 1240.4; DB 6; Length 1254;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1241; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 538 ATTGCTGACCTGAGTGCTGCTAGACCGTCACTGTCACCAACAGGCAATTTGTTGTT 597
Db |||||
QY 7 ATTGCTGACCTGAGTGGCTGCTAGACCGTCACTGTCACCAACAGGCAATTTGTTGTT 66
QY 598 AGCGTTGCTGCTACTGTTGAGGGACGAATCAAGACATTAGTCTTAAATTTTTCGAATC 657
Db |||||
QY 67 AGCGTTGCTGCTACTGTTGAGGGACGAATCAAGACATTAGTCTTAAATTTTTCGAATC 126
QY 658 GATCTAACATCAACGACCTGCTCATGGAGAAAGACAGACGAGGCTTAAGTCCAAATCA 717
Db |||||
QY 127 GATCTAACATCAACGACCTGCTCATGGAGAAAGACAGACGAGGCTTAAGTCCAAATCA 186
QY 718 AAACCAATTTGCTACTGATGATGCGCGGATGTCAATAAATTTGAGAAAGCTGACTACTA 777
Db |||||
QY 187 AAACCAATTTGCTACTGATGATGCGCGGATGTCAATAAATTTGAGAAAGCTGACTACTA 246
QY 778 AAGCTATTCAAGAACAAATTCATCGCTTAACGTCACAGTAAAGCAGCTACTTTGAGTGC 837
Db |||||
QY 247 AAGCTATTCAAGAACAAATTCATCGCTTAACGTCACAGTAAAGCAGCTACTTTGAGTGC 306
QY 838 ATTGATTTTTCGAAGCGATGCAACCAATTTACTGATCGAAACGGCAAGGTCTACTTTGCTGAC 897
Db |||||
QY 307 ATTGATTTTTCGAAGCGATGCAACCAATTTACTGATCGAAACGGCAAGGTCTACTTTGCTGAC 366

QY 898 AAGATGGTTCGGTAACCTTGGCGAGCCCAACCTGTCCAGAAATTTTGTCTAAGCGACAT 957
Db |||||
QY 367 AAGATGGTTCGGTAACCTTGGCGAGCCCAACCTGTCCAGAAATTTTGTCTAAGCGACAT 426
Db |||||
QY 958 GTGCGGTTAGACCATATTAAGAAAAACCAATACAAACCAAGCGAAATCTGTGTGATGTG 1017
Db |||||
QY 427 GTGCGGTTAGACCATATTAAGAAAAACCAATACAAACCAAGCGAAATCTGTGTGATGTG 486
QY 1018 GAATACCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGTCTCAAA 1077
Db |||||
QY 487 GAATACCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGTCTCAAA 546
QY 1078 GATACCTAAGCTATTGAACACACTAGCTATCGGTGACACCATCATCTCAAGAAATTA 1137
Db |||||
QY 547 GATACCTAAGCTATTGAACACACTAGCTATCGGTGACACCATCATCACTCAAGAAATTA 606
QY 1138 GCTCAAGCAAAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTTATGAACGTGAC 1197
Db |||||
QY 607 GCTCAAGCAAAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTTATGAACGTGAC 666
QY 1198 TCCTCAATCGTCACTCATGACATGACATTTTCGGTACGATTTTACCAATGGATCAAGAG 1257
Db |||||
QY 667 TCCTCAATCGTCACTCATGACATGACATTTTCGGTACGATTTTACCAATGGATCAAGAG 726
QY 1258 TTTACTTACCGTGTAAAAATCGGAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 1317
Db |||||
QY 727 TTTACTTACCGTGTAAAAATCGGAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 786
QY 1318 AATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1377
Db |||||
QY 787 AATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 846
QY 1378 GAAAAGCCGTATGATCCCTTTGATCGAGTCACTTGAACCTGTTACCATCAATACGTT 1437
Db |||||
QY 847 GAAAAGCCGTATGATCCCTTTGATCGAGTCACTTGAACCTGTTACCATCAATACGTT 906
QY 1438 GATGTCGATACCAACCAATTTGCTTAAAGTGAGCAGCTCTTAACAGCTAGCGAACGTAAC 1497
Db |||||
QY 907 GATGTCGATACCAACCAATTTGCTTAAAGTGAGCAGCTCTTAACAGCTAGCGAACGTAAC 966
QY 1498 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTCAACAATCTC 1557
Db |||||
QY 967 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTCAACAATCTC 1026
QY 1558 GATGCTTTTGGTATTATGACATATACCTTAACCTGGAAGAGTAGAGGATATACGATGAC 1617
Db |||||
QY 1027 GATGCTTTTGGTATTATGACATATACCTTAACCTGGAAGAGTAGAGGATATACGATGAC 1086
QY 1618 ACCAACCGTATCATACCGCTTTATATGGGCAAGGACCCGAGAGAGATGCTAGCTAT 1677
Db |||||
QY 1087 ACCAACCGTATCATACCGCTTTATATGGGCAAGGACCCGAGAGAGATGCTAGCTAT 1146
QY 1678 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAGAAACGAGAAGTTTACAGCTACCTG 1737
Db |||||
QY 1147 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAGAAACGAGAAGTTTACAGCTACCTG 1206
QY 1738 CTTTATACGGGACCTATACCTGATTAACCCCTTAACGACAAA 1779
Db |||||
QY 1207 CGTTATACGGGACACCTATACCTGATTAACCCCTTAACGACAAA 1248

RESULT 9
ID ABA05547
XX ABA05547 standard; DNA; 8893 BP.
AC ABA05547;
XX ABA05547;
DT 26-FEB-2002 (first entry)
XX
DE Maxadilan-streptokinase fusion protein plasmid pRYB3maxetk.
XX
KW Streptokinase; cerebroprotective; cardiant; gene therapy; fusion;

KW vasodilator; thrombolytic; angina; myocardial infarction; stroke;
KW gene therapy; maxadilan; sand fly; plasmid; ds.
OS Lutzomyia longipalpis.
OS Unidentified.
OS Synthetic.
OS Chimeric.
PN WO200185100-A2.
XX
PD 15-NOV-2001.
XX
PF 10-MAY-2001; 2001WO-US015209.
XX
PR 11-MAY-2000; 2000US-00569920.
XX
PA (GEHO) GEN HOSPITAL CORP.
XX Reddy VB, Lerner B;
XX WPI; 2002-062184/08.
DR
XX
CC The invention relates to a fusion protein or a conjugate comprising a
CC vasodilator polypeptide, or its active fragment, and a thrombolytic
CC polypeptide or its active fragment. The protein is useful for treating a
CC subject suffering from a partially or totally occluded blood vessel,
CC causing unstable angina, acute myocardial infarction or stroke. The
CC polynucleotide encoding the polypeptide is useful in gene therapy. The
CC vasodilator action of the protein allows for the use of lower doses of a
CC thrombolytic while maintaining the clot dissolving effectiveness of the
CC thrombolytic, and the use of lower doses of the thrombolytic reduces
CC associated side effects. The present sequence is the plasmid used for the
CC preparation of Maxadilan-Streptokinase fusion protein. It contains
CC maxadilan cDNA, which encodes a vasodilator peptide produced by the
CC salivary gland of the New World sand fly, and streptokinase cDNA
XX
SQ Sequence 8893 BP; 2366 A; 2106 C; 2246 G; 2175 T; 0 U; 0 Other;

Query Match 69.6%; Score 1240.4; DB 6; Length 8893;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1241; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 538 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCCCAATTGGTTGTT 597
Db |||||
QY 5935 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCCCAATTGGTTGTT 5994
QY 598 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAATC 657
Db |||||
QY 5995 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAATC 6054
QY 658 GATCTAAACATCAGACCTGCTCATGAGGAGAAAGACAGACGAAGGCTTAAGTCCAAATFCA 717
Db |||||
QY 6055 GATCTAAACATCAGACCTGCTCATGAGGAGAAAGACAGACGAAGGCTTAAGTCCAAATFCA 6114
QY 718 AAACATTTGCTACTGATAGTGGCGCGATGTCAATAAATTGAGAAAGCTGACTACTA 777
Db |||||
QY 6115 AAACATTTGCTACTGATAGTGGCGCGATGTCAATAAATTGAGAAAGCTGACTACTA 6174
QY 778 AAGGCTATTCAAGAACCAATTGATCGCTAAAGTCCACAGTAGACGACGACTCTTTGAGGTC 837
Db |||||
QY 6175 AAGGCTATTCAAGAACCAATTGATCGCTAAAGTCCACAGTAGACGACGACTCTTTGAGGTC 6234
QY 838 ATTGATTTTGAAGCGATGCAACCATTTACTGATCGAAACGGCAAGGCTTACTTTGCTGAC 897
Db |||||
QY 6235 ATTGATTTTGAAGCGATGCAACCATTTACTGATCGAAACGGCAAGGCTTACTTTGCTGAC 6294

Db 301 ATTGATTTTGCAGCGATGCAACCAATTACTGATCGAAACGCAAGGCTACTTTGCTGAC 360
 Qy 898 AAGATGTTTGGTAACTTCCGACCCACCTGTCACGAATTTTGTCTAAGCGGACAT 957
 Db 361 AAGATGTTTGGTAACTTCCGACCCACCTGTCACGAATTTTGTCTAAGCGGACAT 420
 Qy 958 GTGCGGTTAGACCATATAAAGAAAAACCAATACAAAACCAAGCGAATCTGTTGATGTG 1017
 Db 421 GTGCGGTTAGACCATATAAAGAAAAACCAATACAAAACCAAGCGAATCTGTTGATGTG 480
 Qy 1018 GAATATACGTGACGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAAA 1077
 Db 481 GAATATACGTGACGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAAA 540
 Qy 1078 GATACCTAAGCTATTGAACACACTAGCTATCGGTGACACCATCATCTCAAGAACTACTA 1137
 Db 541 GATACCTAAGCTATTGAACACACTAGCTATCGGTGACACCATCATCTCAAGAACTACTA 600
 Qy 1138 GCTCAAGCAAAAGCAATTTTAAACAAAACCCAGGCTTATACGATTTTATGAACGTGAC 1197
 Db 601 GCTCAAGCAAAAGCAATTTTAAACAAAACCCAGGCTTATACGATTTTATGAACGTGAC 660
 Qy 1198 TCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 1257
 Db 661 TCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 720
 Qy 1258 TTTACTTACCGTTGTTAAATTCGGGAACAGCTTATAGATCAATAAAAAATCTGGTCTG 1317
 Db 721 TTTACTTACCGTTGTTAAATTCGGGAACAGCTTATAGATCAATAAAAAATCTGGTCTG 780
 Qy 1318 AATGAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1377
 Db 781 AATGAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 840
 Qy 1378 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTTCAACCATCAATACGTT 1437
 Db 841 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTTCAACCATCAATACGTT 900
 Qy 1438 GATGTCGATACCAACGAATGCTTAAAGAGTGAGCAGCTTTAAACAGCTAGCGAACGTAAC 1497
 Db 901 GATGTCGATACCAACGAATGCTTAAAGAGTGAGCAGCTTTAAACAGCTAGCGAACGTAAC 960
 Qy 1498 TTAGACTTCAGAGATTTTACGATCCTCGTGATAAGGCTTAACTTCTTCAACCAATCTC 1557
 Db 961 TTAGACTTCAGAGATTTTACGATCCTCGTGATAAGGCTTAACTTCTTCAACCAATCTC 1020
 Qy 1558 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAAAGTAGAGGATAATCAGATGAC 1617
 Db 1021 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAAAGTAGAGGATAATCAGATGAC 1080
 Qy 1618 ACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAGAGAGAAATGCTAGCTAT 1677
 Db 1081 ACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAGAGAGAAATGCTAGCTAT 1140
 Qy 1678 CATTTAGCCTATGATAAAGATCGTTTATACCGAAGAAAGACGAGAAGTTTACAGCTACCTG 1737
 Db 1141 CATTTAGCCTATGATAAAGATCGTTTATACCGAAGAAAGACGAGAAGTTTACAGCTACCTG 1200
 Qy 1738 CGTTATACAGGACACCTTATACCTGATTAACCCCTTAAACGACAAA 1779
 Db 1201 CGTTATACAGGACACCTTATACCTGATTAACCCCTTAAACGACAAA 1242

RESULT 11

AAX16633
 ID AAX16633 standard; DNA; 1242 BP.
 XX AC AAX16633;
 XX DX
 XX 04-MAY-1999 (first entry)
 DE Streptococcus equisimilis mutant streptokinase K59E encoding DNA.

XX Streptococcus equisimilis H46A; streptokinase; mutant; fibrinolytic;
 KW plasma clot; hydrolysis; haemolytic Streptococcus; plasminogen; plasmin;
 KW serine protease; fibrin; blood clot; thrombolytic;
 KW vascular thromboembolytic symptom; acute myocardial infarction;
 KW fibrinolysis; resistance; ds.
 XX Streptococcus dysgalactiae subsp. equisimilis.
 OS Synthetic.
 OS
 FH Key Location/Qualifiers
 CDS 1..1242
 FT /*tag= a
 FT /transl_except= (pos:40..42,aa:Asn)
 FT /note= "no stop codon given"
 FT
 XX US5876999-A.
 PN 02-MAR-1999.
 PD
 PF 06-DEC-1995; 95US-00568393.
 XX
 PR 06-DEC-1995; 95US-00568393.
 XX
 PA (NASC-) NAT SCI COUNCIL.
 XX
 PI Wu H;
 XX
 DR WPI; 1999-189643/16.
 DR P-PSDB; AAX94665.
 XX
 PT Mutant streptokinase polypeptide - useful as plasmin-resistant
 PT thrombolytic agent.
 XX
 PS Claim 4; Col 11-14; 17pp; English.
 CC The present invention describes a mutant streptokinase (SK) polypeptide
 CC in which at least one of the amino acids in the Pro58-Lys59-Ser60-Lys61
 CC segment of the corresponding native SK is replaced by another amino acid.
 CC The present sequence encodes mutant SK K59E. SK is a secretory protein of
 CC haemolytic Streptococcus able to activate human plasminogen (HPIg) to
 CC plasmin (HPI), which is a serine protease able to catalyse the
 CC hydrolysis of fibrin in blood clots. The SK is useful as a thrombolytic
 CC agent in the treatment of vascular thromboembolytic symptoms such as
 CC acute myocardial infarction. Compared with wild-type SK, the K59E mutant
 CC is more resistant to degradation by human plasmin and is more effective
 CC both in acting as a fibrinolytic agent and in activating human plasminogen
 XX
 SQ Sequence 1242 BP; 421 A; 268 C; 239 G; 314 T; 0 U; 0 Other;
 Query Match 69.3%; Score 1235.6; DB 2; Length 1242;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 1238; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Qy 538 ATTGCTCGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCCAATTGGTTGTT 597
 Db 1 ATTGCTCGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCCAATTGGTTGTT 60
 Qy 598 AGCGTTGCTGTACTGTTGAGGGACCGAATCAAGACATTAGTCTTAAATTTTTTGAATC 657
 Db 61 AGCGTTGCTGTACTGTTGAGGGACCGAATCAAGACATTAGTCTTAAATTTTTTGAATC 120
 Qy 658 GATCTAAATCAACGACCTGCTCATGAGGAGAAAGACAGAGCAAGGCTTAAGTCCAAATCA 717
 Db 121 GATCTAAATCAACGACCTGCTCATGAGGAGAAAGACAGAGCAAGGCTTAAGTCCGGAATCA 180
 Qy 718 AAACCATTTGCTACTGATGAGCGCGATGTGCATATAACTTGGAAAGCTGACTTACTA 777
 Db 181 AAACCATTTGCTACTGATGAGCGCGATGTGCATATAACTTGGAAAGCTGACTTACTA 240
 Qy 778 AAGGCTATTCAAGAACCAATTGATCGCTAAACCTGACAGTAAACGACGACTTCTTTGAGGTC 837
 Db 241 AAGGCTATTCAAGAACCAATTGATCGCTAAACCTGACAGTAAACGACGACTTCTTTGAGGTC 300

QY 838 ATTGATTTTGCAGCGATGCAACATTACTGATCGAAACCGCAAGGCTTACTTTGCTGAC 897
DB |||||
QY 301 ATTGATTTTGCAGCGATGCAACATTACTGATCGAAACCGCAAGGCTTACTTTGCTGAC 360
DB |||||
QY 898 AAGATGGTTCGGTACCTTCCGCGACCAACCTGTCGAAGATTTTGTCAAGCGGACAT 957
DB |||||
QY 361 AAGATGGTTCGGTACCTTCCGCGACCAACCTGTCGAAGATTTTGTCAAGCGGACAT 420
QY 958 GTGCGGTTAGACCATATAAGAAACCAACCAATACAAACCAAGCGAATCTGTTGATGTG 1017
DB |||||
QY 421 GTGCGGTTAGACCATATAAGAAACCAACCAATACAAACCAAGCGAATCTGTTGATGTG 480
QY 1018 GAATATACGTATGATGATTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTCAAA 1077
DB |||||
QY 481 GAATATACGTATGATGATTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTCAAA 540
QY 1078 GATACCTAAGCTATTGAACACACTAGCTATCGGTGACACCATCATCTCAAGAAATTA 1137
DB |||||
QY 541 GATACCTAAGCTATTGAACACACTAGCTATCGGTGACACCATCATCTCAAGAAATTA 600
QY 1138 GCTCAAGCAAAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTATGAACGTGAC 1197
DB |||||
QY 601 GCTCAAGCAAAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTATGAACGTGAC 660
QY 1198 TCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 1257
DB |||||
QY 661 TCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 720
QY 1258 TTTACTACCGTGTAAACATCGGGAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 1317
DB |||||
QY 721 TTTACTACCGTGTAAACATCGGGAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 780
QY 1318 AATGAAGAAATAACCACTGACCTGATCTCTGAGAAATATTACGCTTAAAAAGGG 1377
DB |||||
QY 781 AATGAAGAAATAACCACTGACCTGATCTCTGAGAAATATTACGCTTAAAAAGGG 840
QY 1378 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTTGAAACTGTTCCACCATCAATACGTT 1437
DB |||||
QY 841 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTTGAAACTGTTCCACCATCAATACGTT 900
QY 1438 GATGTCGATACCAACGAATGCTTAAAGAGTGAGCAGCTCTTAAAGCTAGCGAAGCTAAC 1497
DB |||||
QY 901 GATGTCGATACCAACGAATGCTTAAAGAGTGAGCAGCTCTTAAAGCTAGCGAAGCTAAC 960
QY 1498 TTAGACTTCAGAGATTTATACGATCTCTGATGATAGGCTAAACTCTCTACAAATCTC 1557
DB |||||
QY 961 TTAGACTTCAGAGATTTATACGATCTCTGATGATAGGCTAAACTCTCTACAAATCTC 1020
QY 1558 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAAGCTAGAGGATATACGATGAC 1617
DB |||||
QY 1021 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAAGCTAGAGGATATACGATGAC 1080
QY 1618 ACCAAGCGTATCATACCGTTTATATGCGGCAAGCGACCGAAGGAGAGATGCTAGCTAT 1677
DB |||||
QY 1081 ACCAAGCGTATCATACCGTTTATATGCGGCAAGCGACCGAAGGAGAGATGCTAGCTAT 1140
QY 1678 CATTTAGCCTATGATTAAGATCGTTTATACCGAAGAGAAACGAGAGTTTACAGTACCTG 1737
DB |||||
QY 1141 CATTTAGCCTATGATTAAGATCGTTTATACCGAAGAGAAACGAGAGTTTACAGTACCTG 1200
QY 1738 CGTTATACAGGACACCTATACCTGATTAACCCCTAACGAGAAA 1779
DB |||||
QY 1201 CGTTATACAGGACACCTATACCTGATTAACCCCTAACGAGAAA 1242

RESULT 12

AAA37628 standard; DNA; 1327 BP.

XX AAA37628;

XX AC

DT 15-SEP-2003 (revised)

DT 13-OCT-2000 (first entry)
XX Streptokinase-NTR gene.
DE
XX Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
KW plasminogen; human; fibronectin; thrombolytic therapy;
KW cardiovascular disorder; ss.
XX Streptococcus dysgalactiae subsp. equisimilis.
OS
XX
PN EP1024192-A2.
XX
PD 02-AUG-2000.
XX
PF 23-DEC-1999; 99EP-00310541.
XX
XX 24-DEC-1998; 99IN-DE003825.
PR (COUL) CSIR COUNCIL SCI IND RES.
XX
PI Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
PI Yadav M;
XX
DR WPI; 2000-516032/47.
XX
XX Hybrid streptokinase-fibrin binding domain polypeptides useful for
PT thrombolytic therapy comprises a streptokinase fused with fibrin binding
PT domains of human fibronectin.
XX

Example 1; Fig 14; 58pp; English.

This sequence represents a streptokinase-NTR (SK-NTR) gene (where NTR stands for N-terminally repaired with native sequence). The invention relates to a hybrid plasminogen activator (PA) comprises a polypeptide fusion between streptokinase (SK), which are capable of plasminogen (PG) activation, and fibrin binding regions of human fibronectin, which are from fibrin binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA possesses the ability to bind with fibrin independently and also characteristically retains a PG activation ability which becomes evident only after a pronounced duration, or lag, after exposure of the PA to a suitable animal or human PG. The hybrid streptokinase-fibrin binding domain polypeptides are useful in thrombolytic therapy for various kinds of cardiovascular disorders. The hybrids have enhanced fibrin selectivity as well as kinetics of plasminogen activation that are distinct from that of natural streptokinase in being characterised by a temporary delay, or lag of several minutes in the natural rate of the catalytic conversion of plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins can bind tightly with fibrin in blood clots soon after introduction into the vascular system without significantly activating the circulating blood plasminogen to plasmin, thus aiding in the localisation of the plasminogen activation process to the site of pathological thrombus. This overcomes systemic plasminogen activation encountered during clinical use of streptokinase. (Updated on 15-SEP-2003 to standardise OS field)

Sequence 1327 BP; 456 A; 279 C; 250 G; 342 T; 0 U; 0 Other;

Query Match 69.1%; Score 1232.2; DB 3; Length 1327;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1237; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 538 ATTGCTGACCTGAGTGGCTGCTAGACGGTCCATCTGTCAACACAGCCAAATGGTGT 597
DB |||||
QY 83 ATAGCTGGTCTGTAATGGCTACTAGATCGTCTTCTGTAATAACAGCCAAATGGTGT 142
DB |||||
QY 598 AGCGTTCGTGCTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 657
DB |||||
QY 143 AGCGTTCGTGCTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 202
QY 658 GATCTAAACATCACGACCTGCTCTATGAGGAGAAAGACAGACGAGGCTTAAGTCCAAATCA 717
DB |||||
QY 203 GATCTAAACATCACGACCTGCTCTATGAGGAGAAAGACAGACGAGGCTTAAGTCCAAATCA 262
QY 718 AAACATTTGCTACTGATAGTGGCGCGATGTCAATAACTTGAGAAAGCTGACTACTA 777

Db 263 AARCAATTTGCTACTGATAGTGGCGGATGTCATATAAATTCAGAAAGCTGACTTACTA 322
Qy 778 AAGGCTATTCAAGAAACAATTTGATCGCTTAACGTCACAGTAACGACGACTACTTTGAGGTC 837
Db 323 AAGGCTATTCAAGAAACAATTTGATCGCTTAACGTCACAGTAACGACGACTACTTTGAGGTC 382
Qy 838 ATTGATTTTGCAGCGATGCAACATTTACTGATCGAAACGCGCAAGGTCTACTTTGCTGAC 897
Db 383 ATTGATTTTGCAGCGATGCAACATTTACTGATCGAAACGCGCAAGGTCTACTTTGCTGAC 442
Qy 898 AAGATGTTTGGTAACTTTGCGGACCCAACTGTGCAAGAAATTTTGTGTAAGCGGCAT 957
Db 443 AAGATGTTTGGTAACTTTGCGGACCCAACTGTGCAAGAAATTTTGTGTAAGCGGCAT 502
Qy 958 GTGCGGTTTAGACCATATAAAGAAACCAATACCAACCAAGCGAAATCTGTTGATGTG 1017
Db 503 GTGCGGTTTAGACCATATAAAGAAACCAATACCAACCAAGCGAAATCTGTTGATGTG 562
Qy 1018 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTCAAA 1077
Db 563 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTCAAA 622
Qy 1078 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTTACTA 1137
Db 623 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTTACTA 682
Qy 1138 GCTCAAGCAAAAGCAATTTTAAACAAACCCAGGCTATAGGATTTATGAACGTGAC 1197
Db 683 GCTCAAGCAAAAGCAATTTTAAACAAACCCAGGCTATAGGATTTATGAACGTGAC 742
Qy 1198 TCCTCAATCGTCACTCATGACATGACATTTTCGCTACGATTTTACCAATGGATCAAGAG 1257
Db 743 TCCTCAATCGTCACTCATGACATGACATTTTCGCTACGATTTTACCAATGGATCAAGAG 802
Qy 1258 TTTACTTACCGTGTAAAAATCGGAAACAGCTTTATAGGATCAATAAAAAATCTGGTCTG 1317
Db 803 TTTACTTACCGTGTAAAAATCGGAAACAGCTTTATAGGATCAATAAAAAATCTGGTCTG 862
Qy 1318 AATGAAGAAATAAACAACGCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1377
Db 863 AATGAAGAAATAAACAACGCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 922
Qy 1378 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACTGTTCCACATCAATAGTT 1437
Db 923 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACTGTTCCACATCAATAGTT 982
Qy 1438 GATGTCGATACCAACGAAATGCTTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 1497
Db 983 GATGTCGATACCAACGAAATGCTTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 1042
Qy 1498 TTAGACTTCAGAGATTTATACGATCTCTGATGATGAGCTTAACTCTACAAATCTC 1557
Db 1043 TTAGACTTCAGAGATTTATACGATCTCTGATGATGAGCTTAACTCTACAAATCTC 1102
Qy 1558 GATGCTTTTGGTATTAGACTATACCTTAACTGGAAGTAGAGGATTAATCAGATGAC 1617
Db 1103 GATGCTTTTGGTATTAGACTATACCTTAACTGGAAGTAGAGGATTAATCAGATGAC 1162
Qy 1618 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAGGAGAGATGCTAGCTAT 1677
Db 1163 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAGGAGAGATGCTAGCTAT 1222
Qy 1678 CATTTAGCCTATGATAAAGATCGTTTATACCGAAGAGAGAGAGTTTACAGCTACCTG 1737
Db 1223 CATTTAGCCTATGATAAAGATCGTTTATACCGAAGAGAGAGAGTTTACAGCTACCTG 1282
Qy 1738 CGTTTATACGGGACACCTATACCTGATTAACCCCTAAACGAAATAA 1782
Db 1283 CGTTTATACGGGACACCTATACCTGATTAACCCCTAAACGAAATAA 1327

AAT77778

ID AAT77778 standard; cDNA; 2566 BP.

XX AAT77778;

AC AAT77778;

XX 17-OCT-2003 (revised)

DT 01-OCT-1997 (first entry)

XX Coding sequence for plasminogen-binding fragment of Streptokinase.

XX Plasminogen-binding fragment; streptokinase; degradation; MBP;

KW thrombolytic agent; blood clot; bolus; maltose-binding protein; ds.

XX Streptococcus dysgalactiae subsp. equisimilis.

XX W09641883-A1.

XX 27-DEC-1996.

XX 07-JUN-1996; 96WO-US009640.

XX 09-JUN-1995; 95US-0048940.

XX (HARD) HARVARD COLLEGE.

XX Reed GL;

XX WPI; 1997-065469/06.

XX Modified forms of streptokinase resistant to enzymatic cleavage - useful

PT as thrombolytic agents in treating thrombosis and in medical equipment.

XX Example 1; Page 22-23; 65pp; English.

CC This sequence encodes the wild type plasminogen-binding fragment of

CC streptokinase. The protein fragment encoded by this sequence was used in

CC the design of a modified streptokinase has an in vitro degradation rate

CC at least 2 times slower than that of native streptokinase. Compounds

CC containing modified streptokinases are specifically used as thrombolytic

CC agents for dissolving blood clots in vivo in a mammal, preferably at a

CC dose of 20000 U/kg, opt. as a bolus rather than by continuous infusion.

CC (Updated on 17-OCT-2003 to standardise OS field)

XX SQ Sequence 2566 BP; 825 A; 545 C; 469 G; 727 T; 0 U; 0 Other;

Query Match 69.0%; Score 1230.4; DB 2; Length 2566;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 1243; Conservative 0; Mismatches 1; Indels 2; Gaps 1;

Qy 537 TATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAAATTTGGTTGT 596

Db 896 TATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAAATTTGGTTGT 955

Qy 597 TAGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAAT 656

Db 956 TAGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAAT 1015

Qy 657 CGATCTAACATCACGACCTGCTCATGGAGGAAGACAGACGCAAGGCTTTAAGTCCAAATC 716

Db 1016 CGATCTAACATCACGACCTGCTCAT--AGGAACACAGACGCAAGGCTTTAAGTCCAAATC 1073

Qy 717 AAAACCAATTTGCTACTGATAGTGGCGGATGTCAATAACTTGAGAAAGCTGACTTACT 776

Db 1074 AAAACCAATTTGCTACTGATAGTGGCGGATGTCAATAACTTGAGAAAGCTGACTTACT 1133

Qy 777 AAAGGCTATTCAAGAAACAATTTGATCGCTTAACGTCACAGTAACGACGACTTTTGAGGT 836

Db 1134 AAAGGCTATTCAAGAAACAATTTGATCGCTTAACGTCACAGTAACGACGACTTTTGAGGT 1193

Qy 837 CATTTGATTTTGCAGCGATGCAACCATTTACTGTAGTGAAGCGCAAGGCTCTACTTTGCTGA 896

Db 1194 CATTTGATTTTGCAGCGATGCAACCATTTACTGTAGTGAAGCGCAAGGCTCTACTTTGCTGA 1253

QY 897 CAAAGATGGTTCGGTAACTTGGCGGACCCAACTGTCCAGAAATTTTGTCTAAGCGGACA 956
Dd |||||
QY 1254 CAAAGATGGTTCGGTAACTTGGCGGACCCAACTGTCCAGAAATTTTGTCTAAGCGGACA 1313
Dd |||||
QY 957 TGTGCGCGTTAGACCATATAAGAAACCAATATCAAAACCAAGCGGAATCTGTTGATGT 1016
Dd |||||
QY 1314 TGTGCGCGTTAGACCATATAAGAAACCAATATCAAAACCAAGCGGAATCTGTTGATGT 1373
Dd |||||
QY 1017 GGAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTCAAGCAGGTCTCAA 1076
Dd |||||
QY 1374 GGAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTCAAGCAGGTCTCAA 1433
Dd |||||
QY 1077 AGATACCTAGCTATTGAAACACATAGCTATCGGTGACACCATCATCTCAAGAAATTA 1136
Dd |||||
QY 1434 AGATACCTAGCTATTGAAACACATAGCTATCGGTGACACCATCATCTCAAGAAATTA 1493
Dd |||||
QY 1137 AGCTCAAGCAGCAAGCAATTTTAAACAAAACCCAGGCTATACGATTTATGAACGTGA 1196
Dd |||||
QY 1494 AGCTCAAGCAGCAAGCAATTTTAAACAAAACCCAGGCTATACGATTTATGAACGTGA 1553
Dd |||||
QY 1197 CTCCTCAATCGTCACATGACATGACATTTTCCGTACGATTTTACCAATGGATCAAGA 1256
Dd |||||
QY 1554 CTCCTCAATCGTCACATGACATGACATTTTCCGTACGATTTTACCAATGGATCAAGA 1613
Dd |||||
QY 1257 GTTTACTTACCGTGTAAATAATCGGAACCAAGCTTATAGGATCAATAAAAAATCTGTCT 1316
Dd |||||
QY 1614 GTTTACTTACCGTGTAAATAATCGGAACCAAGCTTATAGGATCAATAAAAAATCTGTCT 1673
Dd |||||
QY 1317 GAATGAAGAAATAAACAACCTGACCTGATCTCTGAGAAATATATACGCTCTTAAAAAAG 1376
Dd |||||
QY 1674 GAATGAAGAAATAAACAACCTGACCTGATCTCTGAGAAATATATACGCTCTTAAAAAAG 1733
Dd |||||
QY 1377 GAAAGACCGTATGATCCCTTTCATCGCAGTCACTTGAACCTGTTCAACCATCAATACGT 1436
Dd |||||
QY 1734 GAAAGACCGTATGATCCCTTTCATCGCAGTCACTTGAACCTGTTCAACCATCAATACGT 1793
Dd |||||
QY 1437 TGATGTCGATACCAACGAATTTGCTAAAGAGTGAGCAGCTCTTAAACAGCTAGCGAAGCTAA 1496
Dd |||||
QY 1794 TGATGTCGATACCAACGAATTTGCTAAAGAGTGAGCAGCTCTTAAACAGCTAGCGAAGCTAA 1853
Dd |||||
QY 1497 CTTAGACTTCAGAGATTTATACGATCTCTGTGATAGGCTAAACTACTCTACAACTATCT 1556
Dd |||||
QY 1854 CTTAGACTTCAGAGATTTATACGATCTCTGTGATAGGCTAAACTACTCTACAACTATCT 1913
Dd |||||
QY 1557 CGATGCTTTTGTGATTTATGAGCTATACCTTAACTGGAAGAGTACGATTAATCAAGATGA 1616
Dd |||||
QY 1914 CGATGCTTTTGTGATTTATGAGCTATACCTTAACTGGAAGAGTACGATTAATCAAGATGA 1973
Dd |||||
QY 1617 CACCAACCGTATCATAAACCGTTTATATGGCAAGCGACCGAGGAGAGATGCTAGCTA 1676
Dd |||||
QY 1974 CACCAACCGTATCATAAACCGTTTATATGGCAAGCGACCGAGGAGAGATGCTAGCTA 2033
Dd |||||
QY 1677 TCATTTAGCTTATGATTAAGATCGTTATATCCGAAAGAGAACGAGAGCTTTTACAGCTACCT 1736
Dd |||||
QY 2034 TCATTTAGCTTATGATTAAGATCGTTATATCCGAAAGAGAACGAGAGCTTTTACAGCTACCT 2093
Dd |||||
QY 1737 GCGTTATACAGGACACCTATACCTGATTAACCTTACGACAAATAA 1782
Dd |||||
QY 2094 GCGTTATACAGGACACCTATACCTGATTAACCTTACGACAAATAA 2139
Dd |||||

RESULT 14
ID AAQ12162

XX standard; DNA; 1458 BP.
AC AAQ12162;
XX
XX
DT 25-MAR-2003 (revised)
DT 17-SEP-1991 (first entry)
XX
XX Factor Xa-cleavable hirudin-1EGR-streptokinase gene.
XX
KW Fusion protein; blood clotting; coagulation; fibrinolysis;

KW antithrombotic; thrombolysis; ss.
XX Synthetic.
OS
XX
FH Key Location/Qualifiers
FT CDS 1..1452 /*tag= a
FT mat_peptide 1..195 /*tag= b
FT misc_RNA 196..207 /label= hirudin HV-1
FT /*tag= c
FT /label= linker
FT /note= "encodes factor Xa cleavage site"
FT mat_peptide 208..1449 /*tag= d
FT /label= streptokinase
PN WO9109125-A.
XX 27-JUN-1991.
XX 07-DEC-1989; 89GB-00027722.
XX 07-DEC-1989; 89GB-00027722.
XX 07-DEC-1990; 90WO-GB001911.
PA (BRBI-) BRITISH BIO-TECHNOLOGY LTD.
XX Dawson KM, Hunter MG, Czaplowski LG;
PI WPI: 1991-208151/28.
XX P-PSDB; AAR12885.
XX
XX Fusion protein cleavage by blood clotting enzyme - for prodn. of
XX fractions having greater antithrombotic activity for therapy and
XX prophylaxis.
XX
XX Disclosure; Page 96; 115pp; English.
XX
XX The sequence of the synthetic hirudin HV-1 genes was designed based on
XX the published amino acid sequence (Dodd J., et al FEBS Letters 165 180
XX (1984)). The sequence of streptokinase was obt'd. from PCR amplified
XX chromosomal DNA from S. equisimilis ATCC 10009 or ATCC 9642. The primers
XX used for the PCR were based on the published DNA sequence of S.
XX equisimilis strain H46A (Malke, H., Roe, B., and Ferretti, J.J., Gene 34
XX 357-362 [1985]). The two sequences were used to construct an expression
XX vector in which the hirudin gene is linked to the streptokinase gene via
XX a linking sequence encoding a cleavage site for factor Xa. The factor Xa
XX is present at the site of the target thrombus so the active agents are
XX released specifically at the place where clot formation is occurring. See
XX also AAQ12153-Q12156, AAQ12158-Q12161 and AAQ12490. (Updated on 25-MAR-
XX 2003 to correct PA field.)
XX
XX Sequence 1458 BP; 491 A; 316 C; 290 G; 361 T; 0 U; 0 Other;

Query Match 68.7%; Score 1223.4; DB 2; Length 1458;
Best Local Similarity 98.3%; Pred. No. 0;
Matches 1236; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 526 ACCGATGTTGTTGCTTGGACCTGAGTGGCTGTAGACCTGCATCTGTCAACACAGC 585
Dd |||||
Dd 196 ATCGAAGGTAGAATTTGCTGGACCTGAGTGGCTGTAGACCTGCATCTGTCAACACAGC 255
QY 586 CAATTGTTGTTAGCTTGTCTGTTGCTGTTGAGGGGAGGATCAAGCATTAAGTCTTAAA 645
Dd |||||
Dd 256 CAATTGTTGTTAGCTTGTCTGTTGCTGTTGAGGGGAGGATCAAGCATTAAGTCTTAAA 315
QY 646 TTTTTCGAAATTCGATCTTAACATACGACCTCTCATGAGGAGGAGACAGACGAGGCTTA 705
Dd |||||
Dd 316 TTTTTCGAAATTCGATCTTAACATACGACCTCTCATGAGGAGGAGACAGACGAGGCTTA 375
QY 706 AGTCCAAATCAAAACCAATTTGCTGATGATGGCGGATGTCAATATAAATCTTGAGAAA 765

Db 376 AGTCCAAAACCAAACTTTGCTACTGATAGTGGCGGATGCCACATAAACTTGAATA 435
QY 766 GCTGACTTACTAAAGGCTATTCAAGAACAAATTGATCGCTAAAGTCCACAGTAACGAGC 825
Db 436 GCTGACTTACTAAAGGCTATTCAAGAACAAATTGATCGCTAAAGTCCACAGTAACGAGC 495
QY 826 TACTTTTGAGGTCAATTGATTTTGCAGCGATGCAACCAATTACTGATGCAAAACGCAAGGTC 885
Db 496 TACTTTTGAGGTCAATTGATTTTGCAGCGATGCAACCAATTACTGATGCAAAACGCAAGGTC 555
QY 886 TACTTTTGCTGACAAAGAGTGGTTGCGTAACCTTTGCGGACCCCAACCTGTCGAAGAAATTTTG 945
Db 556 TACTTTTGCTGACAAAGAGTGGTTGCGTAACCTTTGCGGACCCCAACCTGTCGAAGAAATTTTG 615
QY 946 CTAAGCGGACATGTCGGCGTTAGACCATATAAGAAAAACCAATACAAAACCAAGCGAAA 1005
Db 616 CTAAGCGGACATGTCGGCGTTAGACCATATAAGAAAAACCAATACAAAACCAAGCGAAA 675
QY 1006 TCTGTTGATGTGGAATATATCTGACAGTTTACTCCCTTAAACCCCTGATGACGATTTTACA 1065
Db 676 TCTGTTGATGTGGAATATATCTGACAGTTTACTCCCTTAAACCCCTGATGACGATTTTACA 735
QY 1066 CCAGGTCTCAAGATATAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCT 1125
Db 736 CCAGGTCTCAAGATATAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCT 795
QY 1126 CAAGAAATTAAGTCAAGCAACAAGCAATTTTAAACAAAAACCAACCCAGGCTATAGATT 1185
Db 796 CAAGAAATTAAGTCAAGCAACAAGCAATTTTAAACAAAAACCAACCCAGGCTATAGATT 855
QY 1186 TATGAACGTGACTCTCAATCGTCACTCATGATGCAATGATATTTTCCGTACGATTTTACCA 1245
Db 856 TATGAACGTGACTCTCAATCGTCACTCATGATGCAATGATATTTTCCGTACGATTTTACCA 915
QY 1246 ATGGATCAAGAGTTTACTTTACCGTGTAAATAATCGGGAACAAGCTTATAGATCAATAA 1305
Db 916 ATGGATCAAGAGTTTACTTTACCATGTCAAAATTCGGGAACAAGCTTATAGATCAATAA 975
QY 1306 AAATCTGTTGAATGAAGAAATAAACAACACTGATCTCTGAGAAATATTACGTC 1365
Db 976 AAATCTGTTGAATGAAGAAATAAACAACACTGATCTCTGAGAAATATTACGTC 1035
QY 1366 CTTAAAAAGGGAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTTACC 1425
Db 1036 CTTAAAAAGGGAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTTACC 1095
QY 1426 ATCAATACGTGTGATGTGATACCAACGAATTTGCTAAAAAGTGACGAGCTCTTAAACAGCT 1485
Db 1096 ATCAATACGTGTGATGTCAACCAACGAATTTGCTAAAAAGTGACGAGCTCTTAAACAGCT 1155
QY 1486 AGCGAACGTAACTTAGACTTTCAGAGATTTATACGATTCCTCGTGATAGGCTAAACTACTC 1545
Db 1156 AGCGAACGTAACTTAGACTTTCAGAGATTTATACGATTCCTCGTGATAGGCTAAACTACTC 1215
QY 1546 TACAACATCTCGATGCTTTTGGTATTATGGACTATACCTTAACTGGAAGTAGAGAT 1605
Db 1216 TACAACATCTCGATGCTTTTGGTATTATGGACTATACCTTAACTGGAAGTAGAGAT 1275
QY 1606 AATCAGATGACCAACCGTATACACCGTTTATATGGCAAGCGACCCGAAGGAG 1665
Db 1276 AATCAGATGACCAACCGTATACACCGTTTATATGGCAAGCGACCCGAAGGAG 1335
QY 1666 AATGCTAGCTATCAATTAGCTTATGATAAGATTCGTTATACCGAAGAAGCAAGAGTT 1725
Db 1336 AATGCTAGCTATCAATTAGCTTATGATAAGATTCGTTATACCGAAGAAGCAAGAGTT 1395
QY 1726 TACAGCTACCTCGGTTATACAGGACACCTTATACCTGATAACCCCTAACGCAAAATA 1782
Db 1396 TACAGCTACCTCGGTTATACAGGACACCTTATACCTGATAACCCCTAACGCAAAATA 1452

ADM01294

ID ADM01294 standard; DNA; 7057 BP.

XX

AC ADM01294;

XX

DT 01-JUL-2004 (first entry)

XX

DE Plasmid pVTPA-Bstrep DNA sequence.

XX

XX vector; plasmid; artificial intergene region; plant;

KW transplasmatic angiosperm; agronomic property; stress resistance;

XX rbcL gene; ds.

XX

OS Unidentified.

XX

PN WO2004029256-A2.

XX

PD 08-APR-2004.

XX

PF 15-SEP-2003; 2003WO-CU000009.

XX

PF 27-SEP-2002; 2002CU-00000208.

XX

PA (INGG-) CENT ING GENETICA & BIOTECNOLOGIA.

XX

PI Selman-Housein Sosa G, Aguilar Cabeza E, Gonzalez Quintero ADC;

PI Ramos Gonzalez O;

XX

XX WPI; 2004-316131/29.

XX

XX DNA vector for transformation and expression in plasmids, useful e.g. for

PT producing pharmaceutical proteins or improving agronomic properties, has

PT gene inserted in artificial intergene region.

XX

PS Example 10; SEQ ID NO 26; 74pp; Spanish.

XX

XX The invention relates to a DNA vector (A) for stable transformation and

CC expression of genes (I) in plasmids, where (I) is inserted in an

CC artificial intergene region (AIR) formed by combining two 5'-untranslated

CC regions (5'-UTRs) of genes that transcribe in different directions and

CC are derived from plants of different divisions or classes. (A) are used

CC to produce transplasmatic angiosperms that have improved agronomic

CC properties (e.g. resistance to biotic or abiotic stress) or express a

CC very wide range of agricultural, veterinary, pharmaceutical, nutritional

CC or industrial products, e.g. enzymes, vaccinating antigens, cytokines or

CC immunoglobulins. Use of (A) eliminates the need for a transposon for gene

CC insertion; inserted genes do not require promoters and terminators; and

CC the structure of flanking sequences in (A) ensures universal

CC applicability. Also any selection marker in (A) can be eliminated by

CC homologous recombination. (A) provides efficient and stable expression of

CC genes without causing any functional alterations. This sequence

CC represents the plasmid pVTPA-Estrep, an example of the vector of the

XX invention.

XX

SQ Sequence 7057 BP; 2029 A; 1543 C; 1608 G; 1877 T; 0 U; 0 Other;

QY

Query Match 68.6%; Score 1222.6; DB 12; Length 7057;

Best Local Similarity 98.9%; Pred. No. 0;

Matches 1231; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY

538 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACGCAATTCGTTGTT 597

Db

2023 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACGCAATTCGTTGTT 2082

QY

598 AGCGTTCTCGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAATTTTGAATC 657

Db

2083 AGCGTTCTCGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAATTTTGAATC 2142

QY

658 GATCTTAACATCAGACCTGCTCATGGAGGAAAGACAGACGAGGCTTAAGTCCAAATCA 717

Db

2143 GACCTTAACATCAGACCTGCTCATGGAGGAAAGACAGACGAGGCTTAAGTCCAAATCA 2202

QY

718 AAACCATTTCTGCTACTGATAGTGGCGGATGTCAATAACTTGAGAAAGCTGACTACTA 777

Job time : 1078.25 secs

```
Db 2203 AAACCAATTTGCTACTGATAGTGGCGGATGCCACATAAACTTGAANAAGCTGACTTACTA 2262
QY 778 AAGGCTATTCAAGAAACAATTGATCGCTAAAGTTCACAGTAACGACGACTACTTTGAGGTC 837
Db 2263 AAGGCTATTCAAGAAACAATTGATCGCTAAAGTTCACAGTAACGACGACTACTTTGAGGTC 2322
QY 838 ATTGATTTTGCAGAGCGATGCAACCAATTACTGATCGAAACGGCAAGTCTACTTTGCTGAC 897
Db 2323 ATTGATTTTGCAGAGCGATGCAACCAATTACTGATCGAAACGGCAAGTCTACTTTGCTGAC 2382
QY 898 AAGATGGTTTCGGTAACTTTGCCGACCCCACTGTCGAAGAAATTTTGTCTAAGCGGACAT 957
Db 2383 AAGATGGTTTCGGTAACTTTGCCGACCCCACTGTCGAAGAAATTTTGTCTAAGCGGACAT 2442
QY 958 GTGCGGTTTAGACCATATAAGAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 1017
Db 2443 GTGCGGTTTAGACCATATAAGAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 2502
QY 1018 GAATATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCAGGTCTCAAA 1077
Db 2503 GAATATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCAGGTCTCAAA 2562
QY 1078 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCAATCTCAAGAAATTAATA 1137
Db 2563 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCAATCTCAAGAAATTAATA 2622
QY 1138 GCTCAAGACCAAGAGCATTTTAAACAAAACCCAGGCTATACGATTTTATGAACGTGAC 1197
Db 2623 GCTCAAGACCAAGAGCATTTTAAACAAAACCCAGGCTATACGATTTTATGAACGTGAC 2682
QY 1198 TCCTCAATCGTCACTCATGACATGATTTTCGGTACGATTTTACCAATGGATCAAGAG 1257
Db 2683 TCCTCAATCGTCACTCATGACATGATTTTCGGTACGATTTTACCAATGGATCAAGAG 2742
QY 1258 TTTTACTACCGTGTAAATAACGGGAAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 1317
Db 2743 TTTTACTACCGTGTAAATAACGGGAAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 2802
QY 1318 AATGAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 1377
Db 2803 AATGAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 2862
QY 1378 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTTCAACATCAAAATACGTT 1437
Db 2863 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTTCAACATCAAAATACGTT 2922
QY 1438 GATGTCGATACCAACGAATTGCTTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 1497
Db 2923 GATGTCACACCAACGAATTGCTTAAAAAGCGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 2982
QY 1498 TTAGACTTCAGAGATTTATAGATTCCTCGTGATAAGCTTAACTCTCTACAAACAATCTC 1557
Db 2983 TTAGACTTCAGAGATTTATAGATTCCTCGTGATAAGCTTAACTCTCTACAAACAATCTC 3042
QY 1558 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAAGTAGAGGATATACGATGAC 1617
Db 3043 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAAGTAGAGGATATACGATGAC 3102
QY 1618 ACCAACCGGTATCATAAACCGTTTATATGGGCAAGCGACCCGAGGAGAGATGCTAGCTAT 1677
Db 3103 ACCAACCGGTATCATAAACCGTTTATATGGGCAAGCGACCCGAGGAGAGATGCTAGCTAT 3162
QY 1678 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAGAACGAGAGTTTACAGTACCTG 1737
Db 3163 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAGAACGAGAGTTTACAGTACCTG 3222
QY 1738 CGTTATACAGGGACACTATACCTGATACCCCTAACGACAAATAA 1782
Db 3223 CGTTATACAGGGACACTATACCTGATACCCCTAACGACAAATAA 3267
```

Search completed: January 31, 2006, 18:17:51

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2006, 16:24:19 ; Search time 314.619 Seconds
(without alignments)
10068.109 Million cell updates/sec

Title: US-09-940-235-11
Perfect score: 1782
Sequence: 1 tgcgttcacgttcgtcgcg.....ataacctaacgacaaataa 1782

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/prodata/1/ina/1 COMB.seq:*
- 2: /cgn2_6/prodata/1/ina/5 COMB.seq:*
- 3: /cgn2_6/prodata/1/ina/6A COMB.seq:*
- 4: /cgn2_6/prodata/1/ina/6B COMB.seq:*
- 5: /cgn2_6/prodata/1/ina/H COMB.seq:*
- 6: /cgn2_6/prodata/1/ina/PTUS COMB.seq:*
- 7: /cgn2_6/prodata/1/ina/PP COMB.seq:*
- 8: /cgn2_6/prodata/1/ina/RE COMB.seq:*
- 9: /cgn2_6/prodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1245.8	69.9	2385	3 US-09-211-542A-1	Sequence 1, Appli
2	1240.4	69.6	1242	3 US-09-211-542A-5	Sequence 5, Appli
3	1238.8	69.5	1242	2 US-08-568-393B-1	Sequence 1, Appli
4	1235.6	69.3	1242	2 US-08-568-393B-2	Sequence 2, Appli
5	1230.4	69.0	2566	2 US-08-488-940-19	Sequence 19, Appli
6	1223.4	68.7	1458	2 US-07-854-596B-42	Sequence 42, Appli
7	1222.6	68.6	1245	2 US-07-703-778D-1	Sequence 1, Appli
8	1220.4	68.5	1335	2 US-07-854-596B-14	Sequence 14, Appli
9	1220.4	68.5	1512	2 US-07-854-596B-27	Sequence 27, Appli
10	1220.2	68.5	2589	2 US-07-854-596B-34	Sequence 34, Appli
11	1219.4	68.4	1257	2 US-07-854-596B-25	Sequence 25, Appli
12	1219.4	68.4	1317	2 US-07-854-596B-18	Sequence 18, Appli
13	1216.4	68.3	1467	2 US-07-854-596B-46	Sequence 46, Appli
14	1184	66.4	1209	3 US-09-374-038-11	Sequence 11, Appli
15	1184	66.4	1209	3 US-09-374-038-11	Sequence 11, Appli
16	1181.4	66.3	1245	3 US-09-374-038-12	Sequence 12, Appli
17	1181.4	66.3	1245	3 US-09-374-038-12	Sequence 12, Appli
18	1093.6	61.4	1122	3 US-09-374-038-13	Sequence 13, Appli
19	1093.6	61.4	1122	3 US-09-374-038-13	Sequence 13, Appli
20	1093.6	61.4	1158	3 US-09-374-038-14	Sequence 14, Appli
21	1093.6	61.4	1158	3 US-09-374-038-14	Sequence 14, Appli
22	1081	60.7	1119	2 US-07-854-596B-30	Sequence 30, Appli
23	1080.8	60.7	2253	2 US-07-854-596B-39	Sequence 39, Appli
24	1068	59.9	1068	3 US-09-211-542A-11	Sequence 11, Appli

25	1065	59.8	2208	3 US-09-211-542A-3	Sequence 3, Appli
26	829	46.5	1262	9 5240845-3	Patent No. 5240845
27	826	46.4	1242	9 5240845-2	Patent No. 5240845
28	450	25.3	450	3 US-09-211-542A-13	Sequence 13, Appli
29	232.8	13.1	7679	3 US-09-220-132-38	Sequence 38, Appli
30	232.8	13.1	7680	3 US-09-023-655-1289	Sequence 1289, Ap
31	232.8	13.1	7680	6 PCT-US95-09819-6	Sequence 6, Appli
32	232.8	13.1	7803	2 US-08-551-356-1	Sequence 1, Appli
33	232.8	13.1	7803	6 PCT-US93-12687-1	Sequence 1, Appli
34	232.8	13.1	8044	3 US-09-566-921-135	Sequence 135, App
35	229.8	12.9	4811	3 US-09-813-718-13	Sequence 13, Appli
36	229.6	12.9	7705	2 US-08-259-569-16	Sequence 16, Appli
37	229.6	12.9	7705	2 US-08-826-885-16	Sequence 16, Appli
38	229.6	12.9	7705	9 5455158-2	Patent No. 5455158
39	228.8	12.8	4742	3 US-09-813-718-15	Sequence 15, Appli
40	228.4	12.8	4877	3 US-09-813-718-11	Sequence 11, Appli
41	228.4	12.8	5018	3 US-09-813-718-9	Sequence 9, Appli
42	228	12.8	4100	3 US-09-813-718-5	Sequence 5, Appli
43	228	12.8	4682	3 US-09-813-718-3	Sequence 3, Appli
44	228	12.8	4682	3 US-09-813-718-7	Sequence 7, Appli
45	228	12.8	5174	3 US-09-813-718-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-211-542A-1
; Sequence 1, Application US/09211542A
; Patent No. 6210667
; GENERAL INFORMATION:
; APPLICANT: Reed, Guy L.
; TITLE OF INVENTION: BACTERIAL FIBRIN-DEPENDENT PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROMBERG & SUNSTEIN, LLP
; STREET: 125 Summer Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/211,542A
; FILING DATE: 15-December-1998
; CLASSIFICATION: 1653
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/069,497
; FILING DATE: 15-December-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Attorney, Strimpel, Harriet M.
; REGISTRATION NUMBER: 37,008
; REFERENCE/DOCKET NUMBER: 1874/111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)443-9292
; TELEFAX: (617)443-0004
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2385 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2385
US-09-211-542A-1

Query Match 69.9%; Score 1245.8; DB 3; Length 2385;

	Best Local Similarity	99.4%; Matches 1250;	Pred. No. 0; Conservative 0;	Mismatches 7;	Indels 0;	Gaps 0;
Qy	523	TTCACCCATGTTTCGTATTGCTTGGACCTCGAGTGCGCTGTAGACCGTGCATCTGTCAACAAC	582			
Dd	1129	TCCATCGAGGGTAGGATTGCTGGACCTGAGTGCGCTGTAGACCGTGCATCTGTCAACAAC	1188			
Qy	583	AGCCAATTGTTGTTTGAAGTTGCGTTGCTGTTGAGGGACGAATCAAGAATTAGTGTCTT	642			
Dd	1189	AGCCAATTGTTGTTTGAAGTTGCGTTGCTGTTGAGGGACGAATCAAGAATTAGTGTCTT	1248			
Qy	643	AAATTTTTTGGAAATCGAATCTTAACATCAGCACCTGCTCATGGAGGAAGACAGAGCAAGGC	702			
Dd	1249	AAATTTTTTGGAAATCGAATCTTAACATCAGCACCTGCTCATGGAGGAAGACAGAGCAAGGC	1308			
Qy	703	TTAAGTCCAAAATCAAACCATTGCTACTGATAGTGGCGCGATGTGCATAAACATTGAG	762			
Dd	1309	TTAAGTCCAAAATCAAACCATTGCTACTGATAGTGGCGCGATGTGCATAAACATTGAG	1368			
Qy	763	AAAGCTGACTTACTAAAGGCTATTCAAGAAACAATTGATCGCTAAACGTCACAGTAACGAC	822			
Dd	1369	AAAGCTGACTTACTAAAGGCTATTCAAGAAACAATTGATCGCTAAACGTCACAGTAACGAC	1428			
Qy	823	GACTACTTTGAGGTCATTGATTTTGGCAAGCATGCAACATTAATCTGATCGAAACGGCAAG	882			
Dd	1429	GACTACTTTGAGGTCATTGATTTTGGCAAGCATGCAACATTAATCTGATCGAAACGGCAAG	1488			
Qy	883	GTCTACTTTGCTGCAAAAGATGGTTTCGGTAACCTTGGCCGACCCAACTGTGCCAAGAATTT	942			
Dd	1489	GTCTACTTTGCTGCAAAAGATGGTTTCGGTAACCTTGGCCGACCCAACTGTGCCAAGAATTT	1548			
Qy	943	TTGCTAAGCGGACATGTGCGCGTTAGACCATATAAAGAAAAACCAATACAAAAACCAAGCG	1002			
Dd	1549	TTGCTAAGCGGACATGTGCGCGTTAGACCATATAAAGAAAAACCAATACAAAAACCAAGCG	1608			
Qy	1003	AAATCTGTTGATGTGGAATYATCTGTACAGTTTATCTCCCTTAAACCCCTGATGACGATTTT	1062			
Dd	1609	AAATCTGTTGATGTGGAATYATCTGTACAGTTTATCTCCCTTAAACCCCTGATGACGATTTT	1668			
Qy	1063	AGACCAGGTCCTCAAAAGATATCTAAGCTATTGCAAAACACTAGCTATCGGTGACACCATCACA	1122			
Dd	1669	AGACCAGGTCCTCAAAAGATATCTAAGCTATTGCAAAACACTAGCTATCGGTGACACCATCACA	1728			
Qy	1123	TCTCAAGAAATTACTAGCTCAAGCACAAAGCAATTTTAAACAAAAACCAACCCAGGCTATACG	1182			
Dd	1729	TCTCAAGAAATTACTAGCTCAAGCACAAAGCAATTTTAAACAAAAACCAACCCAGGCTATACG	1788			
Qy	1183	ATTATGAACGTGATCTCCTCAATGTGCATCTATGACAAATGACAATTTTCCGTGACATTTTA	1242			
Dd	1789	ATTATGAACGTGATCTCCTCAATGTGCATCTATGACAAATGACAATTTTCCGTGACATTTTA	1848			
Qy	1243	CCAATGATCAAGAGTTTATCTTACCGGTGTTAAAAATCGGGAAACAAGCTTATAGGATCAAT	1302			
Dd	1849	CCAATGATCAAGAGTTTATCTTACCGGTGTTAAAAATCGGGAAACAAGCTTATAGGATCAAT	1908			
Qy	1303	AAAAAATCTGGTCTGAAATGAAGAAATAAACAACAATGACCTGATCTCTGAGAAATATTAC	1362			
Dd	1909	AAAAAATCTGGTCTGAAATGAAGAAATAAACAACAATGACCTGATCTCTGAGAAATATTAC	1968			
Qy	1363	GTCTTTAAAAAAGGGGAAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTTGAACTGTTC	1422			
Dd	1969	GTCTTTAAAAAAGGGGAAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTTGAACTGTTC	2028			
Qy	1423	ACCATCAAAATACGTTGATGTCGATACCAACGAATTCGTAAAGATGAGCAGCTTTAAACA	1482			
Dd	2029	ACCATCAAAATACGTTGATGTCGATACCAACGAATTCGTAAAGATGAGCAGCTTTAAACA	2088			
Qy	1483	GCTAGCGAAAGTAACTTATAGACTTCAGAGATTTTATACGATCCTCGTATAGGCTTAAACTA	1542			
Dd	2089	GCTAGCGAAAGTAACTTATAGACTTCAGAGATTTTATACGATCCTCGTATAGGCTTAAACTA	2148			
Qy	1543	CTCTACAACAATCTCGATGCTTTTGGTATTTATCGACTATATCCTTAACTGGAAAGTAGAG	1602			

```

Db      2149  CTCTACAACAATCTCGATGCTTTTGGTGATTATATGGACTATACCTTAACCTGGAAAGTAGAG  2208
Qy      1603  GATAATCAGCATGACACACCAACCGGTATCATAAACCGTTTTATATGGGCAAGCGACCCGAAAGGA  1662
Db      2209  GATAATCAAGATGACACCAACCGGTATCATAAACCGTTTTATATGGGCAAGCGACCCGAAAGGA  2268
Qy      1663  GAGAACTCTAGCTATCATTTTAGCCTATGATAAAGATCGTTATACCGAAGAAGAACGAGAA  1722
Db      2269  GAGAATGCTAGCTATCATTTTAGCCTATGATAAAGATCGTTATATACCGAAGAAGAACGAGAA  2328
Qy      1723  GTTTACAGCTACCTGCGCTTATACAGGACACCTTATACCTGATACCTTAACCTCAACGACAAA  1779
Db      2329  GTTTACAGCTACCTGCGCTTATACAGGACACCTTATACCTGATACCTTAACCTTAACGACAAA  2385

RESULT 2
US-09-211-542A-5
; Sequence 5, Application US/092111542A
; Patent No. 6210667
; GENERAL INFORMATION:
; APPLICANT: Reed, Guy L.
; TITLE OF INVENTION: BACTERIAL FIBRIN-DEPENDENT PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROMBERG & SUNSTEIN, LLP
; STREET: 125 Summer Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/211,542A
; FILING DATE: 15-December-1998
; CLASSIFICATION: 1653
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/069,497
; FILING DATE: 15-December-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Attorney, Strimpel, Harriet M.
; REGISTRATION NUMBER: 37,008
; REFERENCE/DOCKET NUMBER: 1874/111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)443-9292
; TELEFAX: (617)443-0004
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1242 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1242
US-09-211-542A-5

```

	Query Match	Best Local Similarity	Score 1240.4;	DB 3;	Length 1242;
	Matches 1241;	Conservative	0; Mismatches	1; Indels	0; Gaps
y	538	ATTGCTGGACCTGAGTGGCTGCTAGACCGTGCATCTGTCAACACACGCCAATTGGTTGTT	597		
b	1	ATTGCTGGACCTGAGTGGCTGCTAGACCGTGCATCTGTCAACACACGCCAATTAGTTGTT	60		
y	598	ACGGTGTGCTGCTACTGTTTGAGGGGACGAATCAACACATTAGTCTTAAATTTTTTGAATC	657		
b	61	ACGGTGTGCTGCTACTGTTTGAGGGGACGAATCAACATTAGTCTTAAATTTTTTGAATC	120		

658 GATCTAATCATCAGCCTCTCATGAGGAAAGACAGAGCAAGCGTTAAGTCCAAATCA 717
Db GATCTAATCATCAGCCTCTCATGAGGAAAGACAGAGCAAGCGTTAAGTCCAAATCA 180
718 AAACCAATTTGCTACTGATAGTGGCGGATGTCTACATAAACTTCAGAAAGCTGACTACTA 777
Db AAACCAATTTGCTACTGATAGTGGCGGATGTCTACATAAACTTCAGAAAGCTGACTACTA 240
778 AAGGCTATTCAAGAAACAAATTTGATCGCTTAAGCTCCACAGTAACGACGACTACTTTGAGGTC 837
Db AAGGCTATTCAAGAAACAAATTTGATCGCTTAAGCTCCACAGTAACGACGACTACTTTGAGGTC 300
838 ATTGATTTTGAAGCATGCAACCATTAATGATCGGAAACGGCAAGGCTACTTTGCTGAC 897
Db ATTGATTTTGAAGCATGCAACCATTAATGATCGGAAACGGCAAGGCTACTTTGCTGAC 360
898 AAAGATGTTTGGTAACTTTCGCGCCCAACCTGTCGAAGAAATTTTGTCTAAGCGGCAT 957
Db AAAGATGTTTGGTAACTTTCGCGCCCAACCTGTCGAAGAAATTTTGTCTAAGCGGCAT 420
958 GTGCGGTTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 1017
Db GTGCGGTTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 480
1018 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAA 1077
Db GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAA 540
1078 GATACTAGCTATTGAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTA 1137
Db GATACTAGCTATTGAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTA 600
1138 GCTCAAGCACAAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTTATGAACGTGAC 1197
Db GCTCAAGCACAAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTTATGAACGTGAC 660
1198 TCCTCAATCGTCACTCATGACAAATGACATTTTCGCTACGATTTTACCAATGGAATCAAGAG 1257
Db TCCTCAATCGTCACTCATGACAAATGACATTTTCGCTACGATTTTACCAATGGAATCAAGAG 720
1258 TTTACTTACCGTTTAAATCGGAAACAAAGCTTTATAGATCAATAAAAATCTGCTCTG 1317
Db TTTACTTACCGTTTAAATCGGAAACAAAGCTTTATAGATCAATAAAAATCTGCTCTG 780
1318 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATACGCTCTTAAAAAAGG 1377
Db AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATACGCTCTTAAAAAAGG 840
1378 GAAAGCCGTTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCAACCATCAATACGTT 1437
Db GAAAGCCGTTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCAACCATCAATACGTT 900
1438 GATGTCGATCAACAAGCAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGCAACGTAAC 1497
Db GATGTCGATCAACAAGCAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGCAACGTAAC 960
1498 TTAGACTTCAGAGATTTATACATCTCTGATAGAGCTAAACTCTTCAACAATCTC 1557
Db TTAGACTTCAGAGATTTATACATCTCTGATAGAGCTAAACTCTTCAACAATCTC 1020
1558 GATGCTTTTGGTATTATGAGCTATACCTTTAACTGGAAGAGTAGAGGATAATCAGGATGAC 1617
Db GATGCTTTTGGTATTATGAGCTATACCTTTAACTGGAAGAGTAGAGGATAATCAGGATGAC 1080
1618 ACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAAAGGAGAGATGCTAGCTAT 1677
Db ACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAAAGGAGAGATGCTAGCTAT 1140
1678 CATTTAGCCTATGATAAGATCGTTTATACCGAAGAGAAACGAGAAGTTTACAGCTACCTG 1737
Db CATTTAGCCTATGATAAGATCGTTTATACCGAAGAGAAACGAGAAGTTTACAGCTACCTG 1200
1738 CGTTATACAGGACACCTATACCTGATTAACCCCTTAACGACAA 1779

Db 1201 CGTTATACAGGACACCTATACCTGATATAACCTTAACGACAA 1242
RESULT 3
US-08-568-393B-1
; Sequence 1, Application US/08568393B
; Patent No. 5876999
; GENERAL INFORMATION:
; APPLICANT: Hua-Lin Wu
; APPLICANT: Guey-Yueh Shi
; TITLE OF INVENTION: Preparation of novel streptokinase
; TITLE OF INVENTION: mutants as improved thrombolytic agents
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jeing & Chang
; STREET: Two No. 5876999th Second Street, Suite 290
; CITY: San Jose
; STATE: California
; COUNTRY: USA
; ZIP: 95113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44MB
; MEDIUM TYPE: storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 6.1 on Window 3.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/568,393B
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Chi-Ping Chang
; REGISTRATION NUMBER: 37,798
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (408) 288-8585
; TELEFAX: (408) 288-8386
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1242 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE:
; HYPOTHETICAL: N
; ANTI-SENSE: N
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus equisimilis H46A
; INDIVIDUAL ISOLATE: Malke, H., Roe, B., and Ferretti, J. J.;
; INDIVIDUAL ISOLATE: "Nucleotide sequence of the streptokinase gene from Streptococcus equisimilis H46A" from Gene 34:357-362 (1985).
; CELL TYPE: Streptococcus equisimilis H46A
US-08-568-393B-1
Query Match 69.5%; Score 1238.8; DB 2; Length 1242;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1240; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 538 ATTGCTGGACCTCAGTGGCTGTAGACCGCTCATCTGTCACACACCAATGTTGTT 597
Db 1 ATTGCTGGACCTCAGTGGCTGTAGACCGCTCATCTGTCACACCAATGTTGTT 60
QY 598 AGCGTTGCTGTTACTGTTGAGGGGACGAATCAAGACATTAATCTTAAATTTTGAATC 657
Db 61 AGCGTTGCTGTTACTGTTGAGGGGACGAATCAAGACATTAATCTTAAATTTTGAATC 120
QY 658 GATCTAATCATCAGCCTCTCATGAGGAAAGACAGAGCAAGCGTTAAGTCCAAATCA 717
Db 121 GATCTAATCATCAGCCTCTCATGAGGAAAGACAGAGCAAGCGTTAAGTCCAAATCA 180
QY 718 AAACCAATTTGCTACTGATAGTGGCGGATGTCTACATAAACTTCAGAAAGCTGACTACTA 777

```

Db 181 AAACATTTGCTACTGATAGTGGCGCGATGTCTACATAAACTTGAGAAAGCTGACTTACTA 240
QY 778 AAGGCTATTCAAGAAACAAATTTGATCGCTAAAGTCCACAGTAACGACGACTACTTTGAGTTC 837
Db 241 AAGGCTATTCAAGAAACAAATTTGATCGCTAAAGTCCACAGTAACGACGACTACTTTGAGTTC 300
QY 838 ATTGATTTTGAAGCGATCAACCAATTTACTGATCGAAACGGCAAGGTCTACTTTGCTGAC 897
Db 301 ATTGATTTTGAAGCGATCAACCAATTTACTGATCGAAACGGCAAGGTCTACTTTGCTGAC 360
QY 898 AAGATGGTTGGTAACTTGGGACCCAACTGTCCAGAAATTTTGTCTAAAGCGACAT 957
Db 361 AAGATGGTTGGTAACTTGGGACCCAACTGTCCAGAAATTTTGTCTAAAGCGACAT 420
QY 958 GTGCGGTTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 1017
Db 421 GTGCGGTTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 480
QY 1018 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCAGGTCTCAAA 1077
Db 481 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCAGGTCTCAAA 540
QY 1078 GATATCTAGCTATTGAAACACATAGCTATCGGTGACCAATCAATCTCAAGAAATTA 1137
Db 541 GATATCTAGCTATTGAAACACATAGCTATCGGTGACCAATCAATCTCAAGAAATTA 600
QY 1138 GCTCAAGCACAAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTTATGAACGTGAC 1197
Db 601 GCTCAAGCACAAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTTATGAACGTGAC 660
QY 1198 TCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGAATCAAGAG 1257
Db 661 TCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGAATCAAGAG 720
QY 1258 TTTACTACCGTGTAAATAACCGGAAACAAAGCTTATAGATCAATAAAAAATCTGCTG 1317
Db 721 TTTACTACCGTGTAAATAACCGGAAACAAAGCTTATAGATCAATAAAAAATCTGCTG 780
QY 1318 AATGAAGAAATATAACCACTGACCTGATCTCTGAGAAATATTAACGCTTAAAAAGGG 1377
Db 781 AATGAAGAAATATAACCACTGACCTGATCTCTGAGAAATATTAACGCTTAAAAAGGG 840
QY 1378 GAAAGGCGGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCCACCATCAATACGTT 1437
Db 841 GAAAGGCGGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCCACCATCAATACGTT 900
QY 1438 GATGTCGATACCAACGAATTTGCTAAAGTGAAGCAGCTCTTAAACGCTAGCGAACGTAAC 1497
Db 901 GATGTCGATACCAACGAATTTGCTAAAGTGAAGCAGCTCTTAAACGCTAGCGAACGTAAC 960
QY 1498 TTAGACTTCAGAGATTTATACGATCCCTGATGATAGGCTTAACTACTCTACAAATCTC 1557
Db 961 TTAGACTTCAGAGATTTATACGATCCCTGATGATAGGCTTAACTACTCTACAAATCTC 1020
QY 1558 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAGAGTAGAGGATATACGATGAC 1617
Db 1021 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAGAGTAGAGGATATACGATGAC 1080
QY 1618 ACCAACCGTATCATACCGTTTATATGCGCAAGCGACCCGAGGAGAGAGTGTAGCTAT 1677
Db 1081 ACCAACCGTATCATACCGTTTATATGCGCAAGCGACCCGAGGAGAGAGTGTAGCTAT 1140
QY 1678 CATTTAGCTTATGATAAGATCGTTTATACCGAAGAGAGACGAGAGTTTACGCTACCTG 1737
Db 1141 CATTTAGCTTATGATAAGATCGTTTATACCGAAGAGAGACGAGAGTTTACGCTACCTG 1200
QY 1738 CGTTATACAGGACACCTTATACCTGATTAACCCCTTAACGACAAA 1779
Db 1201 CGTTATACAGGACACCTTATACCTGATTAACCCCTTAACGACAAA 1242

```

RESULT 4

US-08-568-393B-2

```

; Sequence 2, Application US/08568393B
; Patent No. 5876999
; GENERAL INFORMATION:
; APPLICANT: Hua-Lin Wu
; APPLICANT: Guey-yueh Shi
; TITLE OF INVENTION: Preparation of novel streptokinase
; TITLE OF INVENTION: mutants as improved thrombolytic agents
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jeing & Chang
; STREET: Two No. 5876999th Second Street, Suite 290
; CITY: San Jose
; STATE: California
; COUNTRY: USA
; ZIP: 95113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44MB
; MEDIUM TYPE: storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 6.1 on Window 3.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/568,393B
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Chi-Ping Chang
; REGISTRATION NUMBER: 37,798
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (408) 288-8585
; TELEFAX: (408) 288-8386
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1242 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE:
; HYPOTHETICAL: N
; ANTI-SENSE: N
; FEATURE:
; NAME/KEY: SK-K59E
; LOCATION: DNA sequence No. 5876999174 and 175 have been changed
; LOCATION: from AA to GG, and PROTEIN sequence No. 587699959 has been changed
; LOCATION: from Lys to Glu.
; OTHER INFORMATION:
; US-08-568-393B-2

```

```

Query Match 69.3%; Score 1235.6; DB 2; Length 1242;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1238; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 538 ATTGCTGGACCTCAGTGGCTGTAGACCGTCCATCTGTCAACACAGCAATTTGTTGTT 597
Db 1 ATTGCTGGACCTCAGTGGCTGTAGACCGTCCATCTGTCAACACAGCAATTTGTTGTT 60
QY 598 AGCGTTGCTGTTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAATC 657
Db 61 AGCGTTGCTGTTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAATC 120
QY 658 GATCTAACATCAGACCTCTCATGGAGGAAAGACAGACAGGCTTAAGTCCAAATCA 717
Db 121 GATCTAACATCAGACCTCTCATGGAGGAAAGACAGACAGGCTTAAGTCCGGAATCA 180
QY 718 AAACCATTTCTACTGATAGTGGCGGATGTCAATAAACTTGAGAAAGCTGACTTACTA 777
Db 181 AAACCATTTCTACTGATAGTGGCGGATGTCAATAAACTTGAGAAAGCTGACTTACTA 240
QY 778 AAGGCTATTCAAGAAACAAATTTGATCGCTTAACTCCACAGTAACGACGACTACTTTGAGTTC 837
Db 241 AAGGCTATTCAAGAAACAAATTTGATCGCTTAACTCCACAGTAACGACGACTACTTTGAGTTC 300

```


1077 AGATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATTACT 1136
 1434 AGATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATTACT 1493
 1137 AGCTAAGCACAAGCATTTTAAACAAAACCCAGGGCTATACGATTTATGAAACGTGA 1196
 1494 AGCTAAGCACAAGCATTTTAAACAAAACCCAGGGCTATACGATTTATGAAACGTGA 1553
 1197 CTCCTCAATCGTCACATCATGACATGACATTTTCCTACGATTTTACCANTGATCAAGA 1256
 1554 CTCCTCAATCGTCACATCATGACATGACATTTTCCTACGATTTTACCANTGATCAAGA 1613
 1257 GTTTACTTACCGGTGTTAAATAATCGGAAACAGCTTATAGGATCAATAAAAAATCTGGTCT 1316
 1614 GTTTACTTACCGGTGTTAAATAATCGGAAACAGCTTATAGGATCAATAAAAAATCTGGTCT 1673
 1317 GAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATATAGCTCTTAAAAAAGG 1376
 1674 GAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATATAGCTCTTAAAAAAGG 1733
 1377 GAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTACCATCAATACGT 1436
 1734 GAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTACCATCAATACGT 1793
 1437 TGATGTCGATACCAACGAATTGCTAAAGAGTGAGCAGCTCTTAACAGCTAGCGAAGCTAA 1496
 1794 TGATGTCGATACCAACGAATTGCTAAAGAGTGAGCAGCTCTTAACAGCTAGCGAAGCTAA 1853
 1497 CTTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGGTAAACTACTCTACACATCT 1556
 1854 CTTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGGTAAACTACTCTACACATCT 1913
 1557 CGATGCTTTTGGTATTATGAGCTATACCTTAACCTGGAAGAGTACGATTAATCAGATGA 1616
 1914 CGATGCTTTTGGTATTATGAGCTATACCTTAACCTGGAAGAGTACGATTAATCAGATGA 1973
 1617 CACCAACCGTATACATACCGTTTATATGCGAAGCGACCCGAGGAGAGAAATGCTAGCTA 1676
 1974 CACCAACCGTATACATACCGTTTATATGCGAAGCGACCCGAGGAGAGAAATGCTAGCTA 2033
 1677 TCATTAGCTATGATAAAGATCGTTATACCGAAGAGAACGAGAAAGTTTACAGCTACCT 1736
 2034 TCATTAGCTATGATAAAGATCGTTATACCGAAGAGAACGAGAAAGTTTACAGCTACCT 2093
 1737 GCGTTATACAGGAGACCTTATACCTGATACCCCTAACGACAAATAA 1782
 2094 GCGTTATACAGGAGACCTTATACCTGATACCCCTAACGACAAATAA 2139

RESULT 6

US-07-854-596B-42
 ; Sequence 42, Application US/07854596B
 ; Patent No. 5434073

GENERAL INFORMATION:

APPLICANT: Dawson, Keith M
 APPLICANT: Hunter, Michael G
 APPLICANT: Czaplowski, Lloyd G
 TITLE OF INVENTION: Proteins and nucleic acids
 NUMBER OF SEQUENCES: 73
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Dr. John J. McDonnell
 STREET: Ten South Wacker Drive, Suite 3000
 CITY: Chicago
 STATE: IL
 COUNTRY: USA
 ZIP: 60606
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/854,596B

FILING DATE: 03-JUN-1992
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: McDonnell, John J
 REGISTRATION NUMBER: 26,949
 REFERENCE/DOCKET NUMBER: 92,337
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312-715-1000
 TELEFAX: 312-715-1234
 TELEX: 910-221-5317
 INFORMATION FOR SEQ ID NO: 42:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1458 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: 1..1458
 OTHER INFORMATION: /note= "Hirudin-streptokinase
 OTHER INFORMATION: fusion linked by Factor Xa cleavable IEGR"
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..1449
 FEATURE:
 NAME/KEY: mat_peptide
 LOCATION: 1..1449
 US-07-854-596B-42

Query Match 68.7%; Score 1223.4; DB 2; Length 1458;
 Best Local Similarity 98.3%; Pred. No. 0;
 Matches 1236; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

Qy 526 ACCGATGTTTCGTATTTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAACAGC 585
 Db 196 ATCGAAGTAGAATTTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAACAGC 255
 Qy 586 CAATTGTTGTTAGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAA 645
 Db 256 CAATTGTTGTTAGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAA 315
 Qy 646 TTTTTCGAAATCGATCTTAACATCAGACCTGCTCATGGAGGAAAGACAGACGAAAGGCTTA 705
 Db 316 TTTTTCGAAATTTGACCTTAACATCAGACCTGCTCATGGAGGAAAGACAGACGAAAGGCTTA 375
 Qy 706 AGTCCAAAATCAAAAACCAATTTGCTACTGATAGTGGCGGATGTCACTATAAACTTTGAGAAA 765
 Db 376 AGTCCAAAATCAAAAACCAATTTGCTACTGATAGTGGCGGATGCCACATATAAACTTTGAGAAA 435
 Qy 766 GCTGACTTACTAAAGGCTATTTCAGAGAACAAATTGATCGCTAAAGTCCACAGTAAACGACAC 825
 Db 436 GCTGACTTACTAAAGGCTATTTCAGAGAACAAATTGATCGCTAAAGTCCACAGTAAACGACAC 495
 Qy 826 TACTTTGAGGCTCATTTGATTTTGGCAAGGATGTCACACCATTAAGTATCGAAGCGCAAGGTC 885
 Db 496 TACTTTGAGGCTCATTTGATTTTGGCAAGGATGTCACACCATTAAGTATCGAAGCGCAAGGTC 555
 Qy 886 TACTTTGCTGCAAAAGATGTTTCGGTAAACCTTTGCGGACCCCAACCTGTCCAAAGAAATTTTG 945
 Db 556 TACTTTGCTGCAAAAGATGTTTCGGTAAACCTTTGCGGACCCCAACCTGTCCAAAGAAATTTTG 615
 Qy 946 CTAAGCGGACATGTGCGCGTTAGACCATATATAAGAAAAACAATACAAAAACAAGCGGAAA 1005
 Db 616 CTAAGCGGACATGTGCGCGTTAGACCATATATAAGAAAAACAATACAAAAACAAGCGGAAA 675
 Qy 1006 TCTGTTGATGTGGAATATACTGTACAGTTTACTCTCCCTTAACCCCTGATGAGATTTTCA 1065
 Db 676 TCTGTTGATGTGGAATATACTGTACAGTTTACTCTCCCTTAACCCCTGATGAGATTTTCA 735
 Qy 1066 CAGGTCTCAAAAGATACCTTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACATCT 1125
 Db 736 CAGGTCTCAAAAGATACCTTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACATCT 795

1126 CAAGAATTACTAGCTCAAGCACAAGAGCAATTTTAAACAAAACCAACCCAGGCTATACGATT 1185
1186 TATGACGGTACCTCCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCA 1245
856 TATGAACGGTACCTCCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCA 915
1246 ATGGATCAAGAGTTTACTTACCGTGTAAATAATCGGAAACAAGCTTATAGGATCAATAAA 1305
916 ATGGATCAAGAGTTTACTTACCATGTCAAAATCGGGAACAGCTTATGAGATCAATAAA 975
1306 AAATCTGGTCTGAATGAAGAAATAAACAACACACTGACCTGATCTCTGAGAAATATTAGGTC 1365
976 AAATCTGGTCTGAATGAAGAAATAAACAACACACTGACCTGATCTCTGAGAAATATTAGGTC 1035
1366 CTTAAAAAGGGGAAAAGCGGTATGATCCCTTTGATCGGAGTCACTTGAACCTGTTCAAC 1425
1036 CTTAAAAAGGGGAAAAGCGGTATGATCCCTTTGATCGGAGTCACTTGAACCTGTTCAAC 1095
1426 ATCAAAATACCTTCATGTCGATACCAACGAATTCGTAATAAGTGACGAGCTCTTAAACAGCT 1485
1096 ATCAAAATACCTTCATGTCGATACCAACGAATTCGTAATAAGTGACGAGCTCTTAAACAGCT 1155
1486 AGCGAAGCTAATCTAGACTTTCAGAGATTTTATACGATCCTCGTGATAAGGCTTAAACTACTC 1545
1156 AGCGAAGCTAATCTAGACTTTCAGAGATTTTATACGATCCTCGTGATAAGGCTTAAACTACTC 1215
1546 TACAACATCTCGATGCTTTTGGTATATAGGACTATACCTTAACTGGAAGTAGAGAT 1605
1216 TACAACATCTCGATGCTTTTGGTATATAGGACTATACCTTAACTGGAAGTAGAGAT 1275
1606 AATCAGATGACACCAACCGTATCATACCGTTTATATGGCAAGCGACCGGAGGAG 1665
1276 AATCAGATGACACCAACCGTATCATACCGTTTATATGGCAAGCGACCGGAGGAG 1335
1666 AATGCTAGCTATCATTTAGCTATGATATAAGATCGTTTATACCGAAGAAAGCAGAGATT 1725
1336 AATGCTAGCTATCATTTAGCTATGATATAAGATCGTTTATACCGAAGAAAGCAGAGATT 1395
1726 TACAGTACCTCGTTTATACAGGACACCTTATACCTTATACCTTATACCTTATACCTTATAC 1782
1396 TACAGTACCTCGTTTATACAGGACACCTTATACCTTATACCTTATACCTTATACCTTATAC 1452

RESULT 7
US-07-703-778D-1
Sequence 1, Application US/07703778D
Patent No. 5296366
GENERAL INFORMATION:
APPLICANT: Garcia, M.P.E. et al
TITLE OF INVENTION: METHOD FOR THE ISOLATION AND EXPRESSION
OF A GENE WHICH CODES FOR STREPTOKINASE, NUCLEOTIDE SEQUENCE
TITLE OF INVENTION: OBTAINED, RECOMBINANT DNA AND TRANSFORMED MICROORGANISMS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Stanger, Michaelson, Spivak and Wallace, Esq.
STREET: Parkway 109 Office Center, 328 Newman Springs Road,
CITY: Red Bank
STATE: New Jersey
COUNTRY: USA
ZIP: 07701
COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/2" 1.44Mb IBM compatible diskette
COMPUTER: IBM PS/2 Model 80
OPERATING SYSTEM: MS-DOS 5.0
SOFTWARE: Microsoft Word for Windows 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07703,778D
FILING DATE: 19910522
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Michaelson, Peter L.
REGISTRATION NUMBER: 30090
REFERENCE/DOCKET NUMBER: Centro-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908)530-6671
TELEFAX: (908)530-6584
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1245 base pairs
TYPE: NUCLEOTIDE WITH CORRESPONDING PROTEIN
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA
ORIGINAL SOURCE:
ORGANISM: Streptococcus equisimilis from group C of Lanfield
ORGANISM: definition
IMMEDIATE SOURCE: ATCC-9542 strain
FEATURE: from 1 to 1245 bp mature peptide
OTHER INFORMATION:
OTHER INFORMATION: Properties: Streptokinase gene
OTHER INFORMATION: The gene product binds to human plasminogen
OTHER INFORMATION: The gene product is an activator of human plasminogen
US-07-703-778D-1

Query Match 68.6%; Score 1222.6; DB 2; Length 1245;
Best Local Similarity 98.9%; Pred No. 0;
Matches 1231; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
Qy 538 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCAATTTGTTGTT 597
Db 1 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCAATTTGTTGTT 60
Qy 598 AGCGTTGCTGTACTGTTGAGGGGACGAATCAAGACATTAAGTTTAAATTTTGAATC 657
Db 61 AGCGTTGCTGTACTGTTGAGGGGACGAATCAAGACATTAAGTTTAAATTTTGAATC 120
Qy 658 GATCTAATCATCACCACCTGCTCATGGAGGAAGACAGCAAGCTTAAAGTCCAAATCA 717
Db 121 GACCTAATCATCACCACCTGCTCATGGAGGAAGACAGCAAGCTTAAAGTCCAAATCA 180
Qy 718 AAACCATTTGCTACTGATAGTGGCGCGATGTACATAAACTTGAGAAAGCTGACTACTA 777
Db 181 AAACCATTTGCTACTGATAGTGGCGCGATGTACATAAACTTGAGAAAGCTGACTACTA 240
Qy 778 AAGCTATTCAGAAACAATTTGATCGCTAAACGTCACAGTACGACGACTACTTTGAGTC 837
Db 241 AAGCTATTCAGAAACAATTTGATCGCTAAACGTCACAGTACGACGACTACTTTGAGTC 300
Qy 838 ATTGATTTTGCAGCGATGCAACCATTTACTGATCGAAACGGCAAGCTTACTTTGCTGAC 897
Db 301 ATTGATTTTGCAGCGATGCAACCATTTACTGATCGAAACGGCAAGCTTACTTTGCTGAC 360
Qy 898 AAAGATGTTTCGGTAACCTTCCCGACCAACCTGTCAGAGAAATTTTGTAAAGCGGACAT 957
Db 361 AAAGATGTTTCGGTAACCTTCCCGACCAACCTGTCAGAGAAATTTTGTAAAGCGGACAT 420
Qy 958 GTGCGGCTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 1017
Db 421 GTGCGGCTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 480
Qy 1018 GAATATATCTGACAGTTTACTTCCCTTAAACCCCTGATGACGATTTTCAAGCGGCTCAAA 1077
Db 481 GAATATATCTGACAGTTTACTTCCCTTAAACCCCTGATGACGATTTTCAAGCGGCTCAAA 540
Qy 1078 GATACTAAGCTATTGAAACACATAGCTATCGGTGACACCATCATCTCAAGAAATTTACTA 1137
Db 541 GATACTAAGCTATTGAAACACATAGCTATCGGTGACACCATCATCTCAAGAAATTTACTA 600
Qy 1138 GCTCAAGCACAAGCAATTTTAAACCAAAACCAACCCAGGCTATACGATTTTATGAACGTGAC 1197
Db 601 GCTCAAGCACAAGCAATTTTAAACCAAAACCAACCCAGGCTATACGATTTTATGAACGTGAC 660

1198 TCCTCAATCGTCACTCATGACAAATGATGATTTTCCGTCAGATTTTACCAATGGATCAAGAG 1257
 661 TCCTCAATCGTCACTCATGACAAATGATGATTTTCCGTCAGATTTTACCAATGGATCAAGAG 720
 1258 TTTACTTACCGTGTAAATAATCGGGAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 1317
 721 TTTACTTACCATGTCAAAAATCGGGAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 780
 1318 AATGAAGAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1377
 781 AATGAAGAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 840
 1378 GAAAAGCCGTATGATCCCTTTTGATCGCAGTCACTTGAAATCTGTTACCATCAATACGTT 1437
 841 GAAAAGCCGTATGATCCCTTTTGATCGCAGTCACTTGAAATCTGTTACCATCAATACGTT 900
 1438 GATGCGATACCAAGAAATGCTTAAAGATGAGCAGCTTTTAAACAGCTAGCGAAGCTAAC 1497
 901 GATGTCACACCAAGAAATGCTTAAAGATGAGCAGCTTTTAAACAGCTAGCGAAGCTAAC 960
 1498 TTAGACTTCAGAGATTTTACGATCTCTGATAGCTGCTGATAAGCTTAAACTACTCTTAAACATCTC 1557
 961 TTAGACTTCAGAGATTTTACGATCTCTGATAGCTGCTGATAAGCTTAAACTACTCTTAAACATCTC 1020
 1558 GATGCTTTTGGTATTATGACATATACCTTAACTGGAAGATGAGGATTAATCAGATGAC 1617
 1021 GATGCTTTTGGTATTATGACATATACCTTAACTGGAAGATGAGGATTAATCAGATGAC 1080
 1618 ACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGGAAGGAGAGAAATGCTAGCTAT 1677
 1081 ACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGGAAGGAGAGAAATGCTAGCTAT 1140
 1678 CATTTAGCTTATGATAAGATCGTTATACCGAAGAGAACGAGAAATGCTAGCTAT 1737
 1141 CATTTAGCTTATGATAAGATCGTTATACCGAAGAGAACGAGAAATGCTAGCTAT 1200
 1738 CGTTATACAGGACACTTATACCTGATAACCTTAAACGACAAATAA 1782
 1201 CGTTATACAGGACACTTATACCTGATAACCTTAAACGACAAATAA 1245

RESULT 8

US-07-854-596B-14

; Sequence 14, Application US/07854596B

; Patent No. 5434073

; GENERAL INFORMATION:

; APPLICANT: Dawson, Keith M
 ; APPLICANT: Hunter, Michael G
 ; APPLICANT: Czaplowski, Lloyd G
 ; TITLE OF INVENTION: Proteins and nucleic acids

; NUMBER OF SEQUENCES: 73

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Dr. John J. McDonnell
 ; STREET: Ten South Wacker Drive, Suite 3000
 ; CITY: Chicago
 ; STATE: IL
 ; COUNTRY: USA
 ; ZIP: 60606

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/854,596B
 ; FILING DATE: 03-JUN-1992

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: McDonnell, John J
 ; REGISTRATION NUMBER: 26,949
 ; REFERENCE/DOCKET NUMBER: 92,337
 ; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 312-715-1000

; TELEFAX: 312-715-1234
 ; TELEX: 910-221-5317
 ; INFORMATION FOR SEQ ID NO: 14:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1335 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: 1..1335
 ; OTHER INFORMATION: /note= "Streptokinase gene from S.
 ; OTHER INFORMATION: equisimilis"
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 7..1326
 ; FEATURE:
 ; NAME/KEY: mat_peptide
 ; LOCATION: 7..1326
 ; US-07-854-596B-14
 Query Match 68.5%; Score 1220.4; DB 2; Length 1335;
 Best Local Similarity 98.7%; Pred. No. 0;
 Matches 1230; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
 QY 537 TATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAAGCCAAATTTGGTTGT 596
 DB 84 TATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAAGCCAAATTTAGTTGT 143
 QY 597 TAGCGTTGCTGGTACTCTGTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAAT 656
 DB 144 TAGCGTTGCTGGTACTCTGTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAAT 203
 QY 657 CGATCTAAACATCAGACCTGCTCATGAGGAAAGACAGAGCAAGGCTTAAAGTCCAAATC 716
 DB 204 TGACCTAAACATCAGACCTGCTCATGAGGAAAGACAGAGCAAGGCTTAAAGTCCAAATC 263
 QY 717 AAAACCAATTTGCTACTGATAGTGGCGGATGTCAATAAACTTGAGAAAGCTGACTTACT 776
 DB 264 AAAACCAATTTGCTACTGATAGTGGCGGATGTCCACATAAACTTGAAAGAGCTGACTTACT 323
 QY 777 AAAGGCTATTCAAGAACAAATTTGATCGCTAACGCTCCACAGTAAACGACGACTACTTTGAGGT 836
 DB 324 AAAGGCTATTCAAGAACAAATTTGATCGCTAACGCTCCACAGTAAACGACGACTACTTTGAGGT 383
 QY 837 CATTTGATTTTGCAGCGATGCAACCAATTAATGATCGAAACGGCAAGGCTACTTTTGCTGA 896
 DB 384 CATTTGATTTTGCAGCGATGCAACCAATTAATGATCGAAACGGCAAGGCTACTTTTGCTGA 443
 QY 897 CAAAGATGGTTGCGTAACTTTGCGGACCCAACTGTCCAAGAAATTTTGTCTAAGCGGACA 956
 DB 444 CAAAGATGGTTGCGTAACTTTGCGGACCCAACTGTCCAAGAAATTTTGTCTAAGCGGACA 503
 QY 957 TGTGCGGTTAGACCATATAAGGAAACCAATAACAAACCAAGGAAATCTGTTGATGT 1016
 DB 504 TGTGCGGTTAGACCATATAAGGAAACCAATAACAAACCAAGGAAATCTGTTGATGT 563
 QY 1017 GGAATATCTGTACAGTTTACTCTCCCTTAAACCCCTGATGACGATTTCAAGACGAGTCTCAA 1076
 DB 564 GGAATATCTGTACAGTTTACTCTCCCTTAAACCCCTGATGACGATTTCAAGACGAGTCTCAA 623
 QY 1077 AGATATCTAAGCTATTGAAAAACATGATGATCGGTTGACCAACATCATCTCTCAAGAAATTACT 1136
 DB 624 AGATATCTAAGCTATTGAAAAACATGATGATCGGTTGACCAACATCATCTCTCAAGAAATTACT 683
 QY 1137 AGCTCAGGACAAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTATGACGCTGA 1196
 DB 684 AGCTCAGGACAAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTATGACGCTGA 743
 QY 1197 CTCTCAATCGTCACTCATGCAATGACATTTTCCGTCAGGATTTTCAACATGATCAAGA 1256
 DB 744 CTCTCAATCGTCACTCATGCAATGACATTTTCCGTCAGGATTTTCAACATGATCAAGA 803

```
QY 1257 GTTACTTACCGTGTAAATAATCGGAACAAGCTTATAGGATCAATAAAAAATCTGGTCT 1316
Db 804 GTTACTTACCATGTCAAAAATCGGAACAAGCTTATAGGATCAATAAAAAATCTGGTCT 863
QY 1317 GAATGAAGAAATAAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGG 1376
Db 864 GAATGAAGAAATAAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGG 923
QY 1377 GGAAGACCGGTATGATCCCTTTGATCGCAGTCACTTGAAAACGTTTACCACCAATAACGT 1436
Db 924 GGAAGACCGGTATGATCCCTTTGATCGCAGTCACTTGAAAACGTTTACCACCAATAACGT 983
QY 1437 TGATGTGATACCAACGAATGTCTAAAGAGTGAAGAGTCTTAAACAGTACGGAACGTAA 1496
Db 984 TGATGTCAACACCAACGAATGTCTAAAGAGTGAAGAGTCTTAAACAGTACGGAACGTAA 1043
QY 1497 CTTAGACTTCAGAGATTTATACGATCTCTGATAGGCTTAAAGCTTACTCTACACCAATCT 1556
Db 1044 CTTAGACTTCAGAGATTTATACGATCTCTGATAGGCTTAAAGCTTACTCTACACCAATCT 1103
QY 1557 CGATGCTTTTGGTATTATGACTATACCTTAACTGGAAAAGTGAAGATATACAGATGA 1616
Db 1104 CGATGCTTTTGGTATTATGACTATACCTTAACTGGAAAAGTGAAGATATACAGATGA 1163
QY 1617 CACCAACCGGTATCATAAACCGTTTATATGGCAAGCGACCCGAAAGGAGAGATGCTAGCTA 1676
Db 1164 CACCAACCGGTATCATAAACCGTTTATATGGCAAGCGACCCGAAAGGAGAGATGCTAGCTA 1223
QY 1677 TCATTTAGCCTATGATAAGATCGTTATACGGAAGAGACGAGAAAGTTTACAGTACCT 1736
Db 1224 TCATTTAGCCTATGATAAGATCGTTATACGGAAGAGACGAGAAAGTTTACAGTACCT 1283
QY 1737 GCGTTATACAGGACACCTATACCTGATAACCCCTAACGACAAATAA 1782
Db 1284 GCGTTATACAGGACACCTATACCTGATAACCCCTAACGACAAATAA 1329
```

RESULT 9

```
US-07-854-596B-27
; Sequence 27,-Application US/07854596B
; Patent No. 5434073
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith M
; APPLICANT: Hunter, Michael G
; APPLICANT: Czaplowski, Lloyd G
; TITLE OF INVENTION: Proteins and nucleic acids
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. John J. McDonnell
; STREET: Ten South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/854,596B
; FILING DATE: 03-JUN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,337
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1234
; TELEFAX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 27:
```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1512 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..1512
; OTHER INFORMATION: /note= "Streptokinase
; OTHER INFORMATION: fused to a yeast alpha-factor"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 7..1503
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 7..1503
US-07-854-596B-27

Query Match 68.5%; Score 1220.4; DB 2; Length 1512;
Best Local Similarity 98.3%; Pred. No. 0;
Matches 1233; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 529 GATGTTTCGTATGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCCAA 588
Db 253 GATAAAGAAATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCCAA 312
QY 589 TTGTTGTTAGCGTTGCTGTACTGTTGAGGGGAGCAATCAAGACATTAGTCTTAAATTT 648
Db 313 TTAGTTGTTAGCGTTGCTGTACTGTTGAGGGGAGCAATCAAGACATTAGTCTTAAATTT 372
QY 649 TTTGAAAATCGATCTAAACATCAGACCTGCTCATGAGGAGAAAGACAGAGCAAGGCTTAAGT 708
Db 373 TTTGAAATTTGACCTAAACATCAGACCTGCTCATGAGGAGAAAGACAGAGCAAGGCTTAAGT 432
QY 709 CCAAAAATCAAAAACCAATTTGCTACTGATAGTGGCGGATGTCTCATATAAACTTTGAGAAAGCT 768
Db 433 CCAAAAATCAAAAACCAATTTGCTACTGATAGTGGCGGATGTCTCATATAAACTTTGAGAAAGCT 492
QY 769 GACTTACTAAAGGCTATTCAAGAACAAATGTATCGCTAAGCTCAAGTAAACAGCACTAC 828
Db 493 GACTTACTAAAGGCTATTCAAGAACAAATGTATCGCTAAGCTCAAGTAAACAGCACTAC 552
QY 829 TTTGAGGTCATTGATTTTGGCAAGCGATGCAACCACTACTGATCGAAACGGCAAGGCTTAC 888
Db 553 TTTGAGGTCATTGATTTTGGCAAGCGATGCAACCACTACTGATCGAAACGGCAAGGCTTAC 612
QY 889 TTTGCTGACAAAAGATGGTTTCGGTAACTTGGCGGACCCCAACCTGTCCAAAGAAATTTTGCTA 948
Db 613 TTTGCTGACAAAAGATGGTTTCGGTAACTTGGCGGACCCCAACCTGTCCAAAGAAATTTTGCTA 672
QY 949 AGCGGACATGTGCGGTTAGACCATATATAAGAAAAACAATACAAAAACAAAGCGAAATCT 1008
Db 673 AGCGGACATGTGCGGTTAGACCATATATAAGAAAAACAATACAAAAACAAAGCGAAATCT 732
QY 1009 GTTGATGTGGAATATATCTGTACAGTTTACTCTCCCTTAAACCTTGATGACGATTTGAGACA 1068
Db 733 GTTGATGTGGAATATATCTGTACAGTTTACTCTCCCTTAAACCTTGATGACGATTTGAGACA 792
QY 1069 GGTCTCAAAAGATATAAGCTATTGAAAACACATAGCTATCGGTGACACCATCACATCTCAA 1128
Db 793 GGTCTCAAAAGATATAAGCTATTGAAAACACATAGCTATCGGTGACACCATCACATCTCAA 852
QY 1129 GAATTACTAGCTCAAGCAACAAGCATTTTAAACAAAAACCCACCCAGGCTATACGATTTAT 1188
Db 853 GAATTACTAGCTCAAGCAACAAGCATTTTAAACAAAAACCCATCCAGGCTATACGATTTAT 912
QY 1189 GAACGTGATCTCCTCAATCGTCACTCATGACAAATGACATTTTCCGTAACGATTTTACCAATG 1248
Db 913 GAACGTGATCTCCTCAATCGTCACTCATGACAAATGACATTTTCCGTAACGATTTTACCAATG 972
QY 1249 GATCAAGAGTTTACTTACCGTGTAAATAATCGGGAACAAGCTTATAGATCAATAAAAAA 1308
```

973	GATCAAGAGTTTACTTACCATGTCTAAAAATCGGGAAACAGCTTATGAGATCAATAAAAAA	1032
1309	TCTGGTCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATAATTACGTCCTT	1368
1033	TCTGGTCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATAATTACGTCCTT	1092
1369	AAAAAGGGGAAAGCGGTATGATCCCTTTGATCGCAGTCACATTGAAACTGTTTCACCATC	1428
1093	AAAAAGGGGAAAGCGGTATGATCCCTTTGATCGCAGTCACATTGAAACTGTTTCACCATC	1152
1429	AAATACGTTGATGTCGATACCAACGAATTCGTAAAAAGTCAGCAGCTCTTTAAACAGCTAGC	1488
1153	AAATACGTTGATGTCGATACCAACGAATTCGTAAAAAGTCAGCAGCTCTTTAAACAGCTAGC	1212
1489	GAAGCTAACTTAGACTTCAGAGATTATACGATCCTCGTGATAGGCTAAACTACTCTAC	1548
1213	GAACGTAACTTAGACTTCAGAGATTATACGATCCTCGTGATAGGCTAAACTACTCTAC	1272
1549	AACAAATCTCGATGCTTTTGGTATATTGGAATATACCTTTAACTGAAAAGTAGAGGATAAT	1608
1273	AACAAATCTCGATGCTTTTGGTATATTGGAATATACCTTTAACTGAAAAGTAGAGGATAAT	1332
1609	CACGATGACACCAACCGGTATCATACCGTTTTATATGGGCAAGCGACCCGGAAGGAGAGAT	1668
1333	CACGATGACACCAACCGGTATCATACCGTTTTATATGGGCAAGCGACCCGGAAGGAGAGAT	1392
1669	GCTAGCTATCATTTAGCTTATGATAAAGATCGTTTATACCGAAGAAGAACGAGAAGTTTAC	1728
1393	GCTAGCTATCATTTAGCTTATGATAAAGATCGTTTATACCGAAGAAGAACGAGAAGTTTAC	1452
1729	AGCTACCTCGGTTTATACAGGACACCTTATACCTGATAACCTTAACGACAAATAA	1782
1453	AGCTACCTCGGTTTATACAGGACACCTTATACCTGATAACCTTAACGACAAATAA	1506

RESULT 10

US-07-854-596B-34	
; Sequence 34, Application US/07854596B	
; Patent No. 5434073	
; GENERAL INFORMATION:	
; APPLICANT: Dawson, Keith M	
; APPLICANT: Hunter, Michael G	
; APPLICANT: Czaplowski, Lloyd G	
; TITLE OF INVENTION: Proteins and nucleic acids	
; NUMBER OF SEQUENCES: 73	
; CORRESPONDENCE ADDRESS:	
; ADDRESSEE: Dr. John J. McDonnell	
; STREET: Ten South Wacker Drive, Suite 3000	
; CITY: Chicago	
; STATE: IL	
; COUNTRY: USA	
; ZIP: 60606	
; COMPUTER READABLE FORM:	
; MEDIUM TYPE: Floppy disk	
; COMPUTER: IBM PC compatible	
; OPERATING SYSTEM: PC-DOS/MS-DOS	
; SOFTWARE: PatentIn Release #1.0, Version #1.25	
; CURRENT APPLICATION DATA:	
; APPLICATION NUMBER: US/07/854,596B	
; FILING DATE: 03-JUN-1992	
; CLASSIFICATION: 435	
; ATTORNEY/AGENT INFORMATION:	
; NAME: McDonnell, John J	
; REGISTRATION NUMBER: 26,949	
; REFERENCE/DOCKET NUMBER: 52,337	
; TELECOMMUNICATION INFORMATION:	
; TELEPHONE: 312-715-1000	
; TELEFAX: 312-715-1234	
; TELEX: 910-221-5317	
; INFORMATION FOR SEQ ID NO: 34:	
; SEQUENCE CHARACTERISTICS:	
; LENGTH: 2589 base pairs	
; TYPE: nucleic acid	

QY 1314 TCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAA 1373
 Db 2115 TCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAA 2174
 QY 1374 AGGGGAAAGCCGATGATCCCTTTGATCGCAGTCACCTTGAAACTGTTTCCACCAATA 1433
 Db 2175 AGGGGAAAGCCGATGATCCCTTTGATCGCAGTCACCTTGAAACTGTTTCCACCAATA 2234
 QY 1434 CGTTGATGTCGATACCAACGAATTTGCTAAAGAGTGAGCAGCTCTTAACAGCTAGCGAACG 1493
 Db 2235 CGTTGATGTCGATACCAACGAATTTGCTAAAGAGTGAGCAGCTCTTAACAGCTAGCGAACG 2294
 QY 1494 TAACTTAGACTTCAGAGATTTATACGATCTCTGATGATAAGGCTAAACTTACTCTCAACAA 1553
 Db 2295 TAACTTAGACTTCAGAGATTTATACGATCTCTGATGATAAGGCTAAACTTACTCTCAACAA 2354
 QY 1554 TCTCGATGCTTTTGGTATTATGAGTATACCTTAACCTGGAAGAGTAGAGATATACGA 1613
 Db 2355 TCTCGATGCTTTTGGTATTATGAGTATACCTTAACCTGGAAGAGTAGAGATATACGA 2414
 QY 1614 TGACACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAGGAGAGATGCTAG 1673
 Db 2415 TGACACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAGGAGAGATGCTAG 2474
 QY 1674 CTATCATTTAGCCTATGATAAGATCGTTATACCGAAGAGAACGAGAGATTTACAGCTA 1733
 Db 2475 CTATCATTTAGCCTATGATAAGATCGTTATACCGAAGAGAACGAGAGATTTACAGCTA 2534
 QY 1734 CTTGCTTATACAGGGACACTATACCTGATACCTTAACCTTAACGACAAATAA 1782
 Db 2535 CTTGCTTATACAGGGACACTATACCTGATACCTTAACCTTAACGACAAATAA 2583

RESULT 11

US-07-854-596B-25
 ; Sequence 25, Application US/07854596B
 ; Patent No. 5434073
 ; GENERAL INFORMATION:
 ; APPLICANT: Dawson, Keith M
 ; APPLICANT: Hunter, Michael G
 ; APPLICANT: Czaplewski, Lloyd G
 ; TITLE OF INVENTION: Proteins and nucleic acids
 ; NUMBER OF SEQUENCES: 73
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Dr. John J. McDonnell
 ; STREET: Ten South Wacker Drive, Suite 3000
 ; CITY: Chicago
 ; STATE: IL
 ; COUNTRY: USA
 ; ZIP: 60606
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/854,596B
 ; FILING DATE: 03-JUN-1992
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: McDonnell, John J
 ; REGISTRATION NUMBER: 26,949
 ; REFERENCE/DOCKET NUMBER: 92,337
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 312-715-1000
 ; TELEFAX: 312-715-1234
 ; TELEX: 910-221-5317
 ; INFORMATION FOR SEQ ID NO: 25:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1257 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear

; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: 1..1257
 ; OTHER INFORMATION: /note= "Methionyl-streptokinase
 ; OTHER INFORMATION: fusion protein"
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 4..1248
 ; FEATURE:
 ; NAME/KEY: mat_peptide
 ; LOCATION: 4..1248
 ; US-07-854-596B-25

Query Match 68.4%; Score 1219.4; DB 2; Length 1257;
 Best Local Similarity 98.7%; Pred. No. 0;
 Matches 1229; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
 QY 538 ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACCAACAGCCAAATTGGTTGTT 597
 Db 7 ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACCAACAGCCAAATTGGTTGTT 66
 QY 598 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAATTC 657
 Db 67 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAATTC 126
 QY 658 GATCTAACATCAACGACCTGCTCATGGAGGAAAGACAGAGCAGGCTTAAGTCCAAATCA 717
 Db 127 GACCTAACATCAACGACCTGCTCATGGAGGAAAGACAGAGCAGGCTTAAGTCCAAATCA 186
 QY 718 AAACCATTTGCTACTGATAGTGGCGCGATGTCCATAAACTTGAGAAAGCTGACTTACTA 777
 Db 187 AAACCATTTGCTACTGATAGTGGCGCGATGTCCATAAACTTGAGAAAGCTGACTTACTA 246
 QY 778 AAGGCTATTCAAGAACCAATTGATCGCTAAACGTCACAGTAAACGACGACTACTTTGAGGTC 837
 Db 247 AAGGCTATTCAAGAACCAATTGATCGCTAAACGTCACAGTAAACGACGACTACTTTGAGGTC 306
 QY 838 ATTGATTTGCAAGCGATGCAACCAATTACTGATGCGCAACGCGCAAGGCTACTTTGCTGAC 897
 Db 307 ATTGATTTGCAAGCGATGCAACCAATTACTGATGCGCAACGCGCAAGGCTACTTTGCTGAC 366
 QY 898 AAAGATGTTTCGGTAACTTCGCCGACCAACCTGTCGAAGAAATTTTTCGAAGCGGACAT 957
 Db 367 AAAGATGTTTCGGTAACTTCGCCGACCAACCTGTCGAAGAAATTTTTCGAAGCGGACAT 426
 QY 958 GTGCGCGTTAGACCATATAAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 1017
 Db 427 GTGCGCGTTAGACCATATAAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 486
 QY 1018 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCCAGACCGAGTCTCAAA 1077
 Db 487 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCCAGACCGAGTCTCAAA 546
 QY 1078 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACCACTACATCTCAAGAAATTACTA 1137
 Db 547 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACCACTACATCTCAAGAAATTACTA 606
 QY 1138 GCTCAAGCACAAAGCATTTTAAAAAACAACCAACCCAGGCTATACGATTTATGAACGTCAC 1197
 Db 607 GCTCAAGCACAAAGCATTTTAAAAAACAACCAACCCAGGCTATACGATTTATGAACGTCAC 666
 QY 1198 TCCTCAATCGTCACTCATGACCAATGACATTTTCCGTAGCATTTTACCAATGGATCAAGAG 1257
 Db 667 TCCTCAATCGTCACTCATGACCAATGACATTTTCCGTAGCATTTTACCAATGGATCAAGAG 726
 QY 1258 TTTACTTACCGTGTAAAAATTCGGGAAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 1317
 Db 727 TTTACTTACCATGTCAAAAAATTCGGGAAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 786
 QY 1318 AATGAAGAAATAAACAACACTGATCTCTGAGAAATATTAGCTCTTAAAAAGGG 1377
 Db 787 AATGAAGAAATAAACAACACTGATCTCTGAGAAATATTAGCTCTTAAAAAGGG 846

```

QY 1378 GAAAAGCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCCACCAATACGTT 1437
Db 847 GAAAAGCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCCACCAATACGTT 906
QY 1438 GATGTCGATACCAACGAATGCTTAAAGTGAGCAGCTCTTAAACAGTACGGAACGTAAC 1497
Db 907 GATGTCGATACCAACGAATGCTTAAAGTGAGCAGCTCTTAAACAGTACGGAACGTAAC 966
QY 1498 TTAGACTTCAGAGATTTATACGATCTCTCGTGAATAGGCTAACTCTTACACCAATCTC 1557
Db 967 TTAGACTTCAGAGATTTATACGATCTCTCGTGAATAGGCTAACTCTTACACCAATCTC 1026
QY 1558 GATGCTTTTGGTATTATGACTACTACCTTAACTGGAAAAGTACGAGGATAATCACGATGAC 1617
Db 1027 GATGCTTTTGGTATTATGACTACTACCTTAACTGGAAAAGTACGAGGATAATCACGATGAC 1086
QY 1618 ACCAACCGTATCATACCGTTTATATGGGCAACGACCCGGAAGGAGAGATGCTAGCTAT 1677
Db 1087 ACCAACCGTATCATACCGTTTATATGGGCAACGACCCGGAAGGAGAGATGCTAGCTAT 1146
QY 1678 CATTTAGCTTATGATAAGATCGTTTATACCGAAGAGAACGAGAGATTTACAGCTACCTG 1737
Db 1147 CATTTAGCTTATGATAAGATCGTTTATACCGAAGAGAACGAGAGATTTACAGCTACCTG 1206
QY 1738 CGTTATACAGGACACCTTATACCTGATACCTTAAACGACCAATAA 1782
Db 1207 CGTTATACAGGACACCTTATACCTGATACCTTAAACGACCAATAA 1251

```

RESULT 12

```

US-07-854-596B-18,
; Sequence 18, Application US/07854596B
; Patent No. 5434073
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith M
; APPLICANT: Hunter, Michael G
; APPLICANT: Czaplowski, Lloyd G
; TITLE OF INVENTION: Proteins and nucleic acids
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. John J. McDonnell
; STREET: Ten South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/854,596B
; FILING DATE: 03-JUN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,337
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1317 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc_feature

```

```

; LOCATION: 1..1317
; OTHER INFORMATION: /note= "OmpA fused to mature
; streptokinase gene"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 4..1308
; FEATURE:
; NAME/KEY: mat peptide
; LOCATION: 4..1308
; US-07-854-596B-18
Query Match 68.4%; Score 1219.4; DB 2; Length 1317;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 1229; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
QY 538 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCCAAATGGTGT 597
Db 67 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCCAAATGGTGT 126
QY 598 AGCGTTCTGCTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTTTGAATC 657
Db 127 AGCGTTCTGCTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTTTGAAT 186
QY 658 GATCTAACATCACGACCTGCTCATGAGAGAAAGACAGAGCAAGGCTTAAAGTCCAAATCA 717
Db 187 GACCTAACATCACGACCTGCTCATGAGAGAAAGACAGAGCAAGGCTTAAAGTCCAAATCA 246
QY 718 AAACCAATTTGCTACTGATAGTGGCGGATGTCAATAACTTTGAGAAAGCTGACTTACTA 777
Db 247 AAACCAATTTGCTACTGATAGTGGCGGATGTCAATAACTTTGAGAAAGCTGACTTACTA 306
QY 778 AAGGCTATTCAAGAAACAATTTGATCGCTAACGTCCACAGTAAACGACGACTCTTTGAGTC 837
Db 307 AAGGCTATTCAAGAAACAATTTGATCGCTAACGTCCACAGTAAACGACGACTCTTTGAGTC 366
QY 838 ATTGATTTTGAAGCGATGCAACCAATTTACTGATCGAAACCGCAAGGCTCTACTTTGCTGAC 897
Db 367 ATTGATTTTGAAGCGATGCAACCAATTTACTGATCGAAACCGCAAGGCTCTACTTTGCTGAC 426
QY 898 AAAGATGGTTGCGTAAACCTTTGCCGACCCAACTGTGCCAAGAAATTTTGTAAAGCGGACAT 957
Db 427 AAAGATGGTTGCGTAAACCTTTGCCGACCCAACTGTGCCAAGAAATTTTGTAAAGCGGACAT 486
QY 958 GTGCGCGTTAGACCAATATAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 1017
Db 487 GTGCGCGTTAGACCAATATAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 546
QY 1018 GAATATATCTGTACAGTTTACTTCCCTTAAACCTTATGATCGATTTTCAAGACGAGTCTCAAA 1077
Db 547 GAATATATCTGTACAGTTTACTTCCCTTAAACCTTATGATCGATTTTCAAGACGAGTCTCAAA 606
QY 1078 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCAATCACATCTCAAGAAATTA 1137
Db 607 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCAATCACATCTCAAGAAATTA 666
QY 1138 GCTCAAGCACAAAGCAATTTTAAACAAAAACCAAGCGCTATACGATTTATGAACGTCAC 1197
Db 667 GCTCAAGCACAAAGCAATTTTAAACAAAAACCAAGCGCTATACGATTTATGAACGTCAC 726
QY 1198 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGATGATCAAG 1257
Db 727 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGATGATCAAG 786
QY 1258 TTTTACTTACCGTGTAAAAAATCGGGAACAAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 1317
Db 787 TTTTACTTACCGTGTAAAAAATCGGGAACAAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 846
QY 1318 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAAGGG 1377
Db 847 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAAGGG 906
QY 1378 GAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAAACCTGTTCAACCAATACGTT 1437

```


907	GAAAGCGGTATGATCCCTTTGATCGCAGTCACCTTGAACATGTTCCACCATCAAAATACGTT	966
1438	GATGTCGATACCAACGAATTGCTAAAAAGTGAGCAGCTCTTTAAACAGCTAGCGAAACGTAAC	1497
967	GATGTCAAACCAACGAATTGCTAAAAAGCGAGCAGCTCTTTAAACAGCTAGCGAAACGTAAC	1026
1498	TTAGACTTTCAGAGATTTTATACGATCCTCGTGATAAGGCTAAACTACTCTTACAACAATCTC	1557
1027	TTAGACTTTCAGAGATTTTATACGATCCTCGTGATAAGGCTAAACTACTCTTACAACAATCTC	1086
1558	GATGCTTTTGGTATTATGGACTACTACCTTAACTCGGAAGCTAGAGGATATACAGATGAC	1617
1087	GATGCTTTTGGTATTATGGACTACTACCTTAACTCGGAAGCTAGAGGATATACAGATGAC	1146
1618	ACCAACCGTATCATAAACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAAATGCTAGCTAT	1677
1147	ACCAACCGTATCATAAACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAAATGCTAGCTAT	1206
1678	CATTTAGCCTTATGATAAAGATCGTTTATACCGAAGAAGAACGAGAAGTTTACAGCTACCTG	1737
1207	CATTTAGCCTTATGATAAAGATCGTTTATACCGAAGAAGAACGAGAAGTTTACAGCTACCTG	1266
1738	CGTTATACAGGGACACCTATACCTGTGATAACCTTAACGACAAATAA	1782
1267	CGTTATACAGGGACACCTATACCTGTGATAACCTTAACGACAAATAA	1311

RESULT 13

```

US-07-854-596B-46
; Sequence 46, Application US/07854596B
; Patent No. 5434073
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith M
; APPLICANT: Hunter, Michael G
; APPLICANT: Czaplewski, Lloyd G
; TITLE OF INVENTION: Proteins and nucleic acids
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. John J. McDonnell
; STREET: Ten South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/854,596B
; FILING DATE: 03-JUN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,337
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1467 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..1467
; OTHER INFORMATION: /note= "Streptokinase-hirudin
; OTHER INFORMATION: fusion linked by Factor Xa-cleavable IEGF"

```


Db 901 GATGTCACACCAACCAATTTGCTAAAGAGCGAGCAGCTCTTAACAGCTAGCGAACGTAAC 960
 Qy 1498 TTAGACTTCAGAGATTTATACGATCTCGTGATAAGGCTAAATCTCTACACCAATCTC 1557
 Db 961 TTAGACTTCAGAGATTTATACGATCTCGTGATAAGGCTAAATCTCTACACCAATCTC 1020
 Qy 1558 GATGCTTTTGGTATTATGAGCTATACCTTAACTGGGAAAAGTAGAGGATATACGATGAC 1617
 Db 1021 GATGCTTTTGGTATTATGAGCTATACCTTAACTGGGAAAAGTAGAGGATATACGATGAC 1080
 Qy 1618 ACCAACCGTATCATACCGTTTATATGTTGGAAGCGACCGAAGAGAGATGCTAGCTAT 1677
 Db 1081 ACCAACCGTATCATACCGTTTATATGTTGGAAGCGACCGAAGAGAGATGCTAGCTAT 1140
 Qy 1678 CATTTAGCTTCATCAATGATCGTTTATACCGAAGAGAGATGCTAGCTAT 1737
 Db 1141 CATTTAGCTTCATCAATGATCGTTTATACCGAAGAGAGATGCTAGCTAT 1200
 Qy 1738 CGTTATACAGGACACCTATACCTGATTAACCCCTAACGACAAA 1779
 Db 1201 CGTTATACAGGACACCTATACCTGATTAACCCCTAACGACAAA 1242

RESULT 14
 US-09-374-038-11
 ; Sequence 11, Application US/09374038
 ; Patent No. 6309873
 ; GENERAL INFORMATION:
 ; APPLICANT: Madrazo, Isis Del Carmen Torrens
 ; APPLICANT: Garcia, Jose De Jesus De La Fuente
 ; APPLICANT: Ojalvo, Ariana Garcia
 ; APPLICANT: Menendez, Alina Seralena
 ; APPLICANT: Escalona, Elder Pupo
 ; APPLICANT: Masso, Julio Raul Fernandez
 ; APPLICANT: Griego, Martha De Jesus Gonzalez
 ; TITLE OF INVENTION: STREPTOKINASE MUTANTS
 ; FILE REFERENCE: Sequence Listings 1-14 re: 976-5
 ; Patent No. 6309873
 ; CURRENT APPLICATION NUMBER: US/09/374,038
 ; CURRENT FILING DATE: 1999-08-13
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 11
 ; LENGTH: 1209
 ; TYPE: DNA
 ; ORGANISM: Streptococcus equisimilis
 US-09-374-038-11

Query Match 66.4%; Score 1184; DB 3; Length 1209;
 Best Local Similarity 98.8%; Pred. No. 0;
 Matches 1193; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
 Qy 575 TCAACAAACGCAATTTGGTGTAGCGTTGCTGTTAGCGGAGCGAATCAAGACA 634
 Db 2 TGAACAAACGCAATTTAGTGTGTAGCGTTGCTGTTAGCGGAGCGAATCAAGACA 61
 Qy 635 TTAGTCTTAAATTTTGGAAATCGATCTAACATCAAGCCTGCTCATGGAGGAAAGACAG 694
 Db 62 TTAGTCTTAAATTTTGGAAATTTGACCTAACATCAAGCCTGCTCATGGAGGAAAGACAG 121
 Qy 695 AGCAAGGCTTAAGTCCAAATCAAAACCAATTTGCTACTGATAGTGGCGGATGTCACATA 754
 Db 122 AGCAAGGCTTAAGTCCAAATCAAAACCAATTTGCTACTGATAGTGGCGGATGCCACATA 181
 Qy 755 AACTTTGAGAAAGCTGACTTAAAGGCTATTCAAGAAACAATGATCGCTAACGTCACACA 814
 Db 182 AACTTTGAGAAAGCTGACTTAAAGGCTATTCAAGAAACAATGATCGCTAACGTCACACA 241
 Qy 815 GTAAACGACCTACTTTGAGGTCATTTGATTTTGAAGCGATGCAACCATTTACTGATCGAA 874
 Db 242 GTAAACGACCTACTTTGAGGTCATTTGATTTTGAAGCGATGCAACCATTTACTGATCGAA 301

Qy 875 ACGCCAAGGCTACTTTTCTGCAAAAGATGGTTGGTAAACCTTGCCGACCCAACTGTCC 934
 Db 302 ACGCCAAGGCTACTTTTCTGCAAAAGATGGTTGGTAAACCTTGCCGACCCAACTGTCC 361
 Qy 935 AAGAAATTTTGTCTAAGCGGACATGTGCGGTTAGACCATATATAAGAAAAACCAATACAAA 994
 Db 362 AAGAAATTTTGTCTAAGCGGACATGTGCGGTTAGACCATATATAAGAAAAACCAATACAAA 421
 Qy 995 ACCAAGCGAAATCTGTTGATGTGGAATATCTGTACAGATTTTACTCCCTTAAACCCCTGATG 1054
 Db 422 ATCAAGCGAAATCTGTTGATGTGGAATATCTGTACAGATTTTACTCCCTTAAACCCCTGATG 481
 Qy 1055 ACGATTTTCAGACAGGCTCTCAAGATATAAGCTATTGAAAAACAATAGCTATCGGTGACA 1114
 Db 482 ACGATTTTCAGACAGGCTCTCAAGATATAAGCTATTGAAAAACAATAGCTATCGGTGACA 541
 Qy 1115 CCAATCATCTCAAGAAATTAAGCTCAAGCAAAAGCAATTTTAAACCAACCCAG 1174
 Db 542 CCAATCATCTCAAGAAATTAAGCTCAAGCAAAAGCAATTTTAAACCAACCCAG 601
 Qy 1175 GCTATACGATTTATGAACGTCCTCAATCGTCACTCATGACAATGACATTTTCCGTA 1234
 Db 602 GCTATACGATTTATGAACGTCCTCAATCGTCACTCATGACAATGACATTTTCCGTA 661
 Qy 1235 CGATTTTACCAATGGATCAAGAGTTTACTTACCGTGTAAAAATCGGGAAACAAGCTTATA 1294
 Db 662 CGATTTTACCAATGGATCAAGAGTTTACTTACCATGTCAAAATCGGGAAACAAGCTTATG 721
 Qy 1295 GGATCAATAAAAAATCTGGTCTGAATGAAGAAATAAAACAACACTGATCTCTGAGA 1354
 Db 722 AGATCAATAAAAAATCTGGTCTGAATGAAGAAATAAAACAACACTGATCTCTGAGA 781
 Qy 1355 AATATTACGCTCTTAAAAAGGGGAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGA 1414
 Db 782 AATATTACGCTCTTAAAAAGGGGAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGA 841
 Qy 1415 AACTGTTTCAACCATCAAAATAGCTGATCGATACCAACGAATTCCTAAAAAGTCAAGCAGC 1474
 Db 842 AACTGTTTCAACCATCAAAATAGCTGATCGATACCAACGAATTCCTAAAAAGTCAAGCAGC 901
 Qy 1475 TCTTAAACAGCTAGCGAACGTAACCTTAGACTTCAGAGATTTTATACGATCCTCGTGAAGG 1534
 Db 902 TCTTAAACAGCTAGCGAACGTAACCTTAGACTTCAGAGATTTTATACGATCCTCGTGAAGG 961
 Qy 1535 CTAAACTACTCTACAAACATCTCGATGCTTTGTTGTTATGGAATATACCTTAACCTGAA 1594
 Db 962 CTAAACTACTCTACAAACATCTCGATGCTTTGTTGTTATGGAATATACCTTAACCTGAA 1021
 Qy 1595 AAGTAGAGGATAATCACGATGACCAACCGTATCATAAACCGTTTATATGGCAAGCGAC 1654
 Db 1022 AAGTAGAGGATAATCACGATGACCAACCGTATCATAAACCGTTTATATGGCAAGCGAC 1081
 Qy 1655 CCGAAGGAGAGAAATGCTAGCTATCATTTAGCTTATGATAAGATCGTTTATACCGAAGAG 1714
 Db 1082 CCGAAGGAGAGAAATGCTAGCTATCATTTAGCTTATGATAAGATCGTTTATACCGAAGAG 1141
 Qy 1715 AACGAGAGTTTACAGCTACTCGGTTATACAGGACACCTTACCTGATACCTTAACG 1774
 Db 1142 AACGAGAGTTTACAGCTACTCGGTTATACAGGACACCTTATACCTGATACCTTAACG 1201
 Qy 1775 ACAATAA 1782
 Db 1202 ACAATAA 1209

RESULT 15
 US-09-658-179-11
 ; Sequence 11, Application US/09658179
 ; Patent No. 6413759
 ; GENERAL INFORMATION:
 ; APPLICANT: Madrazo, Isis Del Carmen Torrens
 ; APPLICANT: Garcia, Jose De Jesus De La Fuente
 ; APPLICANT: Ojalvo, Ariana Garcia

```

; APPLICANT: Menendez, Alina Seralena
; APPLICANT: Escalona, Elder Pupo
; APPLICANT: Masso, Julio Raul Fernandez
; APPLICANT: Griego, Martha De Jesus Gonzalez
; TITLE OF INVENTION: STREPTOKINASE MUTANTS
; FILE REFERENCE: Sequence Listings 1-14 re: 976-S-5
; Patent No. 6413759
; CURRENT APPLICATION NUMBER: US/09/658,179
; CURRENT FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 1209
; TYPE: DNA
; ORGANISM: Streptococcus equisimilis
US-09-658-179-11

```

Query Match	66.4%	Score 1184	DB 3	Length 1209
Best Local Similarity	98.8%	Pred. No. 0		
Matches 1193	Conservative 0	Mismatches 15	Indels 0	Gaps 0
QY	575	TCAACACAGCCAAATCGTGTGTAGCGTTCTGCTGTACTGTGTAGGGGACGGAATCAAGACA	634	
Db	2	TGAACACAGCCAAATAGTTGTAGCGTGTGCTGTACTGTGTAGGGGACGGAATCAAGACA	61	
QY	635	TTAGTCTTAAATTTTTTGGAAATCGATCTAAATCAGCAGCTGCTCATGGAGGAAAGACAG	694	
Db	62	TTAGTCTTAAATTTTTTGGAAATTGACCTTAACATCAGCAGCTGCTCATGGAGGAAAGACAG	121	
QY	695	AGCAAGCTTAAAGTCCAAATCAAAACGATTTGCTACTGATAGTGGCGCGGATGTCAATA	754	
Db	122	AGCAAGCTTAAAGTCCAAATCAAAACGATTTGCTACTGATAGTGGCGCGGATGTCAATA	181	
QY	755	AACCTGAGAAAGCTGACTTACTAAAGGCTATTCAAGAACAAATTCATCGCTAAAGCTCAACA	814	
Db	182	AACCTGAAAGGCTGACTTACTAAAGGCTATTCAAGAACAAATTCATCGCTAAAGCTCAACA	241	
QY	815	GTAACGACGACTACTTTTGAAGGTCAATGATTTTGAAGCGATGCAACCAATTAATCTGATCGAA	874	
Db	242	GTAACGACGACTACTTTTGAAGGTCAATGATTTTGAAGCGATGCAACCAATTAATCTGATCGAA	301	
QY	875	ACGGCAGGTCCTACTTTGCTGACAAAGATGGTTGCGTAACTCTTGCCGACCAACCTGTCC	934	
Db	302	ACGGCAGGTCCTACTTTGCTGACAAAGATGGTTGCGTAACTCTTGCCGACCAACCTGTCC	361	
QY	935	AAGAAATTTTGTAAAGCGGACATGTGCGCGTTAGACCATATAAGAAAAAACCAATACAAA	994	
Db	362	AAGAAATTTTGTAAAGCGGACATGTGCGCGTTAGACCATATAAGAAAAAACCAATACAAA	421	
QY	995	ACCAAGGAAATCTGTGTGATGTGGAATATATCTGTATACAGTTTATCTCCCTTAAACCGCTGATG	1054	
Db	422	ATCAAGCGAAATCTGTGTGATGTGGAATATATCTGTATACAGTTTATCTCCCTTAAACCGCTGATG	481	
QY	1055	ACGATTTTCAGACAGGTCCTCAAAGATACTAAGCTTATTGAAAAACACTAGCTATCGGTGACA	1114	
Db	482	ACGATTTTCAGACAGGTCCTCAAAGATACTAAGCTTATTGAAAAACACTAGCTATCGGTGACA	541	
QY	1115	CCATCATCTCTCAAGAAATTAAGTCTCAAGCACAAGCAATTTTAAACAAAAACCAACCCAG	1174	
Db	542	CCATCATCTCTCAAGAAATTAAGTCTCAAGCACAAGCAATTTTAAACAAAAACCAACCCAG	601	
QY	1175	GCTATAGGATTTATGAACGTGACTCTCTCAATCTGCTACATGACATGACATTTTCCGTA	1234	
Db	602	GCTATAGGATTTATGAACGTGACTCTCTCAATCTGCTACATGACATTTTCCGTA	661	
QY	1235	CGATTTTACCAATGGATCAAGAGTTTACTTTACCGTGTTTAAAAATCGGGAAACAGCTTATA	1294	
Db	662	CGATTTTACCAATGGATCAAGAGTTTACTTTACCAATGTCAAAATTCGGGAACAGCTTATG	721	
QY	1295	GGATCAATAAAAATCTGGTCTGAATGAAGAAATAAAACACATGACCTGATCTCTGAGA	1354	
Db	722	AGATCAATAAAAATCTGGTCTGAATGAAGAAATAAAACACATGACCTGATCTCTGAGA	781	

Search completed: February 1, 2006, 12:43:43
Job time : 315.619 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2006, 17:06:30 ; Search time 1472.42 Seconds
(without alignments)
10008.068 Million cell updates/sec

Title: US-09-940-235-11
Perfect score: 1782
Sequence: 1 tcgcttcacgttcgctcg...ataacctaacgacaataa 1782

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_NA_Main:
1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:
3: /cgn2_6/prodata/1/pubpna/US09A_PUBCOMB.seq:
4: /cgn2_6/prodata/1/pubpna/US09B_PUBCOMB.seq:
5: /cgn2_6/prodata/1/pubpna/US10A_PUBCOMB.seq:
6: /cgn2_6/prodata/1/pubpna/US10B_PUBCOMB.seq:
7: /cgn2_6/prodata/1/pubpna/US10C_PUBCOMB.seq:
8: /cgn2_6/prodata/1/pubpna/US10D_PUBCOMB.seq:
9: /cgn2_6/prodata/1/pubpna/US10E_PUBCOMB.seq:
10: /cgn2_6/prodata/1/pubpna/US11_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1782	100.0	1782	3	US-09-940-235-11
2	1684	94.5	2096	3	US-09-940-235-12
3	1245	69.9	1377	3	US-09-940-235-5
4	1243.4	69.8	1245	3	US-09-940-235-1
5	1232.2	69.1	1327	3	US-09-940-235-6
6	1150.2	64.5	1661	3	US-09-940-235-10
7	1134.2	63.6	1541	3	US-09-940-235-9
8	1081.2	60.7	1323	8	US-10-474-792-657
9	232.8	13.1	2127	6	US-10-210-120-49
10	232.8	13.1	2127	9	US-10-956-157-4288
11	232.8	13.1	2127	9	US-10-909-035-49
12	232.8	13.1	2443	7	US-10-741-601-70
13	232.8	13.1	2443	8	US-10-741-600-238
14	232.8	13.1	2488	7	US-10-741-601-75
15	232.8	13.1	2488	8	US-10-741-600-244
16	232.8	13.1	4295	6	US-10-144-194A-51
17	232.8	13.1	4295	8	US-10-491-566-51
18	232.8	13.1	6510	7	US-10-741-601-72
19	232.8	13.1	6510	8	US-10-741-600-241
20	232.8	13.1	6988	7	US-10-236-392-1
21	232.8	13.1	7361	7	US-10-236-392-3
22	232.8	13.1	7677	9	US-10-956-157-4995
23	232.8	13.1	7679	9	US-10-831-704-38

24	232.8	13.1	7680	3	US-09-964-824A-574	Sequence 574, Appl
25	232.8	13.1	7680	5	US-10-171-311-63	Sequence 63, Appl
26	232.8	13.1	7680	6	US-10-236-031B-69	Sequence 69, Appl
27	232.8	13.1	7680	7	US-10-374-979-75	Sequence 75, Appl
28	232.8	13.1	7680	6	US-10-182-936A-75	Sequence 75, Appl
29	232.8	13.1	7680	7	US-10-641-643-1289	Sequence 1289, Ap
30	232.8	13.1	7680	7	US-10-717-597-222	Sequence 222, App
31	232.8	13.1	7680	8	US-10-788-792-79	Sequence 79, App
32	232.8	13.1	7680	8	US-10-477-238A-654	Sequence 654, App
33	232.8	13.1	7680	8	US-10-680-287A-654	Sequence 654, App
34	232.8	13.1	7680	8	US-10-278-698-88	Sequence 88, Appl
35	232.8	13.1	7680	8	US-10-278-698-603	Sequence 603, App
36	232.8	13.1	7680	9	US-10-843-641A-5877	Sequence 5877, Ap
37	232.8	13.1	7680	9	US-10-477-173-654	Sequence 654, App
38	232.8	13.1	7680	9	US-10-852-335A-52	Sequence 52, Appl
39	232.8	13.1	7795	5	US-10-084-817-2	Sequence 2, Appli
40	232.8	13.1	7823	7	US-10-741-601-77	Sequence 77, Appl
41	232.8	13.1	7823	8	US-10-741-600-245	Sequence 245, App
42	232.8	13.1	7848	7	US-10-741-601-78	Sequence 78, Appl
43	232.8	13.1	7848	8	US-10-741-600-246	Sequence 246, App
44	232.8	13.1	7867	5	US-10-098-841-6	Sequence 6, Appli
45	232.8	13.1	7935	7	US-10-741-601-74	Sequence 74, Appl

ALIGNMENTS

RESULT 1
US-09-940-235-11
; Sequence 11, Application US/09940235
; Publication No. US20030059921A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammar
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-005002
; CURRENT APPLICATION NUMBER: US/09/940,235
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 1782
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hybrid cassette
US-09-940-235-11

Query Match 100.0%; Score 1782; DB 3; Length 1782;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1782; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	TCGCTTCACGCTTCGCTCGGTATCGGTGATTCATTCTGCTAACAGTAAGGCAACCCCGC	60
Db	1	TCGCTTCACGCTTCGCTCGGTATCGGTGATTCATTCTGCTAACAGTAAGGCAACCCCGC	60
Qy	61	CAGCCTAGCCGGGTCTTCAACAGCAGGAGCAGCATATGCGCACCCTGGCGCAGGACCCCA	120
Db	61	CAGCCTAGCCGGGTCTTCAACAGCAGGAGCAGCATATGCGCACCCTGGCGCAGGACCCCA	120
Qy	121	ACGCTGCCGAGATCTCGATCCCGGAAATTAATACACTCATATAGGGAGACACAAC	180

```
Db 121 |||||TCATGCAATGACATTTTCGATGAGTATGAGGAGACCAAC 180
Qy 181 |||||ACTTACCGTGTAAATAATCGGAAACAGCTTATAGGATCAATAAAAATCTGGTCTGAAT 240
Db 181 |||||ACTTACCGTGTAAATAATCGGAAACAGCTTATAGGATCAATAAAAATCTGGTCTGAAT 240
Qy 241 |||||GAAGAAATAACCAACACTGACCTGATCTCTGAGAAATATTAACGTCCTTTAAAAAAGGGGAA 1380
Db 241 |||||GAAGAAATAACCAACACTGACCTGATCTCTGAGAAATATTAACGTCCTTTAAAAAAGGGGAA 1380
Qy 301 |||||AAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTTCAACCATCAAAATACGTTGAT 1440
Db 301 |||||AAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTTCAACCATCAAAATACGTTGAT 1440
Qy 361 |||||GTGATACCAACGAATTTGCTAAAGTGAGCAGCTCTTAAACAGCTAGCGAAGCTAACTTA 1500
Db 361 |||||GTGATACCAACGAATTTGCTAAAGTGAGCAGCTCTTAAACAGCTAGCGAAGCTAACTTA 1500
Qy 421 |||||GACCTTACAGAGATTTATACGATCCTCGTGATAAGGCTAAACTTCTTACAACAATCTCGAT 1560
Db 421 |||||GACCTTACAGAGATTTATACGATCCTCGTGATAAGGCTAAACTTCTTACAACAATCTCGAT 1560
Qy 481 |||||GCTTTTGGTATTTATGAGACTATACCTTAACTGGAAGAGTAGAGGATAATCACGATGACACC 1620
Db 481 |||||GCTTTTGGTATTTATGAGACTATACCTTAACTGGAAGAGTAGAGGATAATCACGATGACACC 1620
Qy 541 |||||AACCGTATCATACCGTTTATATGGCAAGCGACCGAAGAGAGAGATGCTAGCTATCAT 1680
Db 541 |||||AACCGTATCATACCGTTTATATGGCAAGCGACCGAAGAGAGAGATGCTAGCTATCAT 1680
Qy 601 |||||TTAGCTTATGATAAAGATCGTTTATACCGAAGAGAAACGAGAAAGTTTACAGCTACTCTGCT 1740
Db 601 |||||TTAGCTTATGATAAAGATCGTTTATACCGAAGAGAAACGAGAAAGTTTACAGCTACTCTGCT 1740
Qy 661 |||||TATACAGGACACCTTATACCTGATTAACCTTAACGACAAATAA 1782
Db 661 |||||TATACAGGACACCTTATACCTGATTAACCTTAACGACAAATAA 1782
Qy 721 |||||CAATTTGCTATGATGCGCGATGTGCATATAAATCTGAGAAAGCTGACTACTAAAG 780
Db 721 |||||CAATTTGCTATGATGCGCGATGTGCATATAAATCTGAGAAAGCTGACTACTACTAAAG 780
Qy 781 |||||GCTATTCAAGAACTTGTATGCTTAACGCTCCACAGTAAACGACACTACTTTGAGGTCATT 840
Db 781 |||||GCTATTCAAGAACTTGTATGCTTAACGCTCCACAGTAAACGACACTACTTTGAGGTCATT 840
Qy 841 |||||GATTTTGCAGGATGCAACCATTTACTGATCGAAACGGCAAGGTCTACTTTGCTGACAAA 900
Db 841 |||||GATTTTGCAGGATGCAACCATTTACTGATCGAAACGGCAAGGTCTACTTTGCTGACAAA 900
Qy 901 |||||GATGTTTGGTAACTTGTGCGGACCCCACTGTCCAAAGATTTTGTGTAAGCGACATGTG 960
Db 901 |||||GATGTTTGGTAACTTGTGCGGACCCCACTGTCCAAAGATTTTGTGTAAGCGACATGTG 960
Qy 961 |||||CCGTTAGCCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGGAA 1020
Db 961 |||||CCGTTAGCCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGGAA 1020
Qy 1021 |||||TATACCTGACAGTTTACTCCCTTAAACCCCTGATGACGATTTTACAGCAGGTCTCAAGAT 1080
Db 1021 |||||TATACCTGACAGTTTACTCCCTTAAACCCCTGATGACGATTTTACAGCAGGTCTCAAGAT 1080
Qy 1081 |||||ACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACTCTCAAGAAATTTACTAGCT 1140
Db 1081 |||||ACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACTCTCAAGAAATTTACTAGCT 1140
Qy 1141 |||||CAAGCACAAGCATTTTAAACAAAACCCAGGCTATACGATTTTATGAACGTGACTCC 1200
Db 1141 |||||CAAGCACAAGCATTTTAAACAAAACCCAGGCTATACGATTTTATGAACGTGACTCC 1200
Qy 1201 |||||TCAATCGTCACTCATGCAATGACATTTTCGTTACGATTTTACCAATGGATCAAGAGTTT 1260
```

```
Db 1201 |||||TCATGCAATGACATTTTCGATGAGTATGAGGAGTATGAGGAGTATGAGGAGTATGAGGAGTAT 1260
Qy 1261 |||||ACTTACCGTGTAAATAATCGGAAACAGCTTATAGGATCAATAAAAATCTGGTCTGAAT 1320
Db 1261 |||||ACTTACCGTGTAAATAATCGGAAACAGCTTATAGGATCAATAAAAATCTGGTCTGAAT 1320
Qy 1321 |||||GAAGAAATAACCAACACTGACCTGATCTCTGAGAAATATTAACGTCCTTTAAAAAAGGGGAA 1380
Db 1321 |||||GAAGAAATAACCAACACTGACCTGATCTCTGAGAAATATTAACGTCCTTTAAAAAAGGGGAA 1380
Qy 1381 |||||AAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTTCAACCATCAAAATACGTTGAT 1440
Db 1381 |||||AAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTTCAACCATCAAAATACGTTGAT 1440
Qy 1441 |||||GTGATACCAACGAATTTGCTAAAGTGAGCAGCTCTTAAACAGCTAGCGAAGCTAACTTA 1500
Db 1441 |||||GTGATACCAACGAATTTGCTAAAGTGAGCAGCTCTTAAACAGCTAGCGAAGCTAACTTA 1500
Qy 1501 |||||GACCTTACAGAGATTTATACGATCCTCGTGATAAGGCTAAACTTCTTACAACAATCTCGAT 1560
Db 1501 |||||GACCTTACAGAGATTTATACGATCCTCGTGATAAGGCTAAACTTCTTACAACAATCTCGAT 1560
Qy 1561 |||||GCTTTTGGTATTTATGAGACTATACCTTAACTGGAAGAGTAGAGGATAATCACGATGACACC 1620
Db 1561 |||||GCTTTTGGTATTTATGAGACTATACCTTAACTGGAAGAGTAGAGGATAATCACGATGACACC 1620
Qy 1621 |||||AACCGTATCATACCGTTTATATGGCAAGCGACCGAAGAGAGAGATGCTAGCTATCAT 1680
Db 1621 |||||AACCGTATCATACCGTTTATATGGCAAGCGACCGAAGAGAGAGATGCTAGCTATCAT 1680
Qy 1681 |||||TTAGCTTATGATAAAGATCGTTTATACCGAAGAGAAACGAGAAAGTTTACAGCTACTCTGCT 1740
Db 1681 |||||TTAGCTTATGATAAAGATCGTTTATACCGAAGAGAAACGAGAAAGTTTACAGCTACTCTGCT 1740
Qy 1741 |||||TATACAGGACACCTTATACCTGATTAACCTTAACGACAAATAA 1782
Db 1741 |||||TATACAGGACACCTTATACCTGATTAACCTTAACGACAAATAA 1782
```

RESULT 2

```
US-09-940-235-12
; Sequence 12, Application US/09940235
; Publication No. US2003005921A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammara
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/09/940,235
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2096
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hybrid cassette
US-09-940-235-12
```

Query Match 94.5%; Score 1684; DB 3; Length 2096;

Best Local Similarity 99.7%; Pred. No. 0;			
Matches 1687; Conservative 0; Mismatches 5; Indels 0; Gaps 0;			
Qy	1	TGGCTTCACGTTTCGCTCGCTATCGGTGATTCATTCGTGTAACACAGTAAGCAACCCCGC	60
Db	51	TGGCTTCACGTTTCGCTCGCTATCGGTGATTCATTCGTGTAACACAGTAAGCAACCCCGC	110
Qy	61	CAGCTTAGCGGGTCTCTAACGACAGGAGACGATCATGCGCACCCCGTGCGCAGACCCA	120
Db	111	CAGCTTAGCGGGTCTCTAACGACAGGAGACGATCATGCGCACCCCGTGCGCAGACCCA	170
Qy	121	ACGCTGCCGAGATCTCGATCCCGGAAATTAATACGACTCATATAGGAGACCAAC	180
Db	171	ACGCTGCCGAGATCTCGATCCCGGAAATTAATACGACTCATATAGGAGACCAAC	230
Qy	181	GGTTTCCCTCTAGAAATAATTTTGTAACTTTAAGAAAGAGATATACCATGTTGCAAGC	240
Db	231	GGTTTCCCTCTAGAAATAATTTTGTAACTTTAAGAAAGAGATATACCATGTTGCAAGC	290
Qy	241	ACAACAGATGTACCATAGCTGAGAGTGTGTTTGATCATGCTGCGGACTTCCTATGT	300
Db	291	ACAACAGATGTACCATAGCTGAGAGTGTGTTTGATCATGCTGCGGACTTCCTATGT	350
Qy	301	GGTGGAGAAAACGTGGAGAGCGGACGATCATCTTGCACTTCTAGAAATAGATGC	360
Db	351	GGTGGAGAAAACGTGGAGAGCGGACGATCATCTTGCACTTCTAGAAATAGATGC	410
Qy	361	AACGATCAGSACACAAAGGACATCTATAGAAATGGAGACACCTGGACAGAGGATAAT	420
Db	411	AACGATCAGSACACAAAGGACATCTATAGAAATGGAGACACCTGGACAGAGGATAAT	470
Qy	421	CGAGGAAACCTGCTCCAGTGTCATCTGCACAGGCAACCGCCGAGGAGAGTGGAAGTGAG	480
Db	471	CGAGGAAACCTGCTCCAGTGTCATCTGCACAGGCAACCGCCGAGGAGAGTGGAAGTGAG	530
Qy	481	AGGCACACCTCTGTGCGAGACCAATCGAGCGGATCTGGCCCTTCAACCGATGTCGTAAT	540
Db	531	AGGCACACCTCTGTGCGAGACCAATCGAGCGGATCTGGCCCTTCAACCGATGTCGTAAT	590
Qy	541	GCTGGACCTGAGTGGCTGTAGACCTGCATCTGTCAACAGCCAAATGTTGTTAGC	600
Db	591	GCTGGACCTGAGTGGCTGTAGACCTGCATCTGTCAACAGCCAAATGTTGTTAGC	650
Qy	601	GTTCGTGTTACTGTTTCAGGGGAGCAATCAAGACATGATCTTAAATTTTGTGAATCGAT	660
Db	651	GTTCGTGTTACTGTTTCAGGGGAGCAATCAAGACATGATCTTAAATTTTGTGAATCGAT	710
Qy	661	CTAACATCACGACCTGCTCATGAGGAAAGACAGACGAGGCTTAAGTCCAAATCAAAA	720
Db	711	CTAACATCACGACCTGCTCATGAGGAAAGACAGACGAGGCTTAAGTCCAAATCAAAA	770
Qy	721	CAATTTGCTACTGATAGTGGCGGATGTCAATAAACTTGAGAAAGCTGACTTAAG	780
Db	771	CAATTTGCTACTGATAGTGGCGGATGTCAATAAACTTGAGAAAGCTGACTTAAG	830
Qy	781	GCTATTCAAGAACAAATTGATCGCTACGTCACAGTAAACGACACTACTTTGAGTCAAT	840
Db	831	GCTATTCAAGAACAAATTGATCGCTACGTCACAGTAAACGACACTACTTTGAGTCAAT	890
Qy	841	GATTTTGCAGGATGCAACCATTTACTGATCGAAACGGCAAGGCTTACTTTGTCACAAA	900
Db	891	GATTTTGCAGGATGCAACCATTTACTGATCGAAACGGCAAGGCTTACTTTGTCACAAA	950
Qy	901	GATGTTTCGGTAACCTTTCGCGACCCCACTGTCCTCAAGAAATTTTGTAAAGGACATGTG	960
Db	951	GATGTTTCGGTAACCTTTCGCGACCCCACTGTCCTCAAGAAATTTTGTAAAGGACATGTG	1010
Qy	961	CGCGTTAGACCATATAAGAAACCAATACAAAACAGCGAAATCTGTTGATGTGGA	1020
Db	1011	CGCGTTAGACCATATAAGAAACCAATACAAAACAGCGAAATCTGTTGATGTGGA	1070
Qy	1021	TATACGTACGTTTACTCCCTTAACCTTAAACCTGATGAGATTTTCAGACGAGTCTCAAGAT	1080

Db	1071	TATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCGAGTCTCAAGAT	1130
Qy	1081	ACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATTTACTAGCT	1140
Db	1131	ACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATTTACTAGCT	1190
Qy	1141	CAAGCACAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTTATGAACGTGACTCC	1200
Db	1191	CAAGCACAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTTATGAACGTGACTCC	1250
Qy	1201	TCAATCGTCACTCATGACAATGACATTTTCGTAGATTTTACCAATGATCAAGAGTTT	1260
Db	1251	TCAATCGTCACTCATGACAATGACATTTTCGTAGATTTTACCAATGATCAAGAGTTT	1310
Qy	1261	ACTTACCGTGTAAATAATCGGGAACAAGCTTATAGATCAATAAAAAATCTGGTCTGAAT	1320
Db	1311	ACTTACCGTGTAAATAATCGGGAACAAGCTTATAGATCAATAAAAAATCTGGTCTGAAT	1370
Qy	1321	GAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAGCTCTTAAAAAAGGGAA	1380
Db	1371	GAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAGCTCTTAAAAAAGGGAA	1430
Qy	1381	AAGCGGTATGATCCCTTTTGGATCGGAGTCACTTTGAAAACCTGTTCAACATCAATCTGAT	1440
Db	1431	AAGCGGTATGATCCCTTTTGGATCGGAGTCACTTTGAAAACCTGTTCAACATCAATCTGAT	1490
Qy	1441	GTGATACCAAGCAATTCCTAAAGTGAGAGCTCTTAAACAGCTAGGAGAGTAACTTA	1500
Db	1491	GTGATACCAAGCAATTCCTAAAGTGAGAGCTCTTAAACAGCTAGGAGAGTAACTTA	1550
Qy	1501	GACTTCAGAGATTTTATACGATCTCTGATTAAGGCTAAACTACTCTACAAATCTCGAT	1560
Db	1551	GACTTCAGAGATTTTATACGATCTCTGATTAAGGCTAAACTACTCTACAAATCTCGAT	1610
Qy	1561	GCTTTTGGTATTATGGACTATACCTTAACTGAAAAGTAGAGGATAATCAGATGACACC	1620
Db	1611	GCTTTTGGTATTATGGACTATACCTTAACTGAAAAGTAGAGGATAATCAGATGACACC	1670
Qy	1621	AACGTTATCATACCGTTTATATGGGCAAGGACCCCGAGAGAGAGATGCTAGCTATCAT	1680
Db	1671	AACGTTATCATACCGTTTATATGGGCAAGGACCCCGAGAGAGAGATGCTAGCTATCAT	1730
Qy	1681	TTAGCTTATGAT 1692	
Db	1731	TTAGCTTGTGT 1742	

RESULT 3
US-09-940-235-5
; Sequence 5, Application US/09940235
; Publication No. US20030059921A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammar
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/09/940,235
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 05/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5

```

; LENGTH: 1377
; TYPE: DNA
; ORGANISM: Streptococcus equisimilis
US-09-940-235-5

Query Match      69.9%; Score 1245; DB 3; Length 1377;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 538 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAAATGGTTGTT 597
Db |||
Qy 133 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAAATGGTTGTT 192
Db |||
Qy 598 ACGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAATC 657
Db |||
Qy 193 ACGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAATC 252
Db |||
Qy 658 GATCTAAACATCAGACCTGTCTCATGAGGAGAAAGACAGACGAGGCTTAAAGTCCCAATCA 717
Db |||
Qy 253 GATCTAAACATCAGACCTGTCTCATGAGGAGAAAGACAGACGAGGCTTAAAGTCCCAATCA 312
Db |||
Qy 718 AAACATTTGCTACTGATAGTGGCGGATGTCAATAAATCTTGAGAAAGCTGACTACTA 777
Db |||
Qy 313 AAACATTTGCTACTGATAGTGGCGGATGTCAATAAATCTTGAGAAAGCTGACTACTA 372
Db |||
Qy 778 AAGGCTATTCAAGAACAAATTTGATCGCTTAAGTCCACAGTAACGACGACTACTTTGAGTTC 837
Db |||
Qy 373 AAGGCTATTCAAGAACAAATTTGATCGCTTAAGTCCACAGTAACGACGACTACTTTGAGTTC 432
Db |||
Qy 838 ATTGATTTTTCGAGCGATGCAACATTAATCTGATCGAAACCGGCAAGGCTCTACTTTGCTGAC 897
Db |||
Qy 433 ATTGATTTTTCGAGCGATGCAACATTAATCTGATCGAAACCGGCAAGGCTCTACTTTGCTGAC 492
Db |||
Qy 898 AAGATGTTTTCGTTAACTTTGCGGACCCAACTGTCTCAAGAAATTTTGTCTAAACGCGACAT 957
Db |||
Qy 493 AAGATGTTTTCGTTAACTTTGCGGACCCAACTGTCTCAAGAAATTTTGTCTAAACGCGACAT 552
Db |||
Qy 958 GTGCGGTTAGACCATATAAAGGAAACCAATACAAACCAAGGGAATCTGTTGATGTG 1017
Db |||
Qy 553 GTGCGGTTAGACCATATAAAGGAAACCAATACAAACCAAGGGAATCTGTTGATGTG 612
Db |||
Qy 1018 GAATATACGTGACGTTTACTCCCTTAAACCCCTGATGACGATTTACAGCAGGCTCTCAAA 1077
Db |||
Qy 613 GAATATACGTGACGTTTACTCCCTTAAACCCCTGATGACGATTTACAGCAGGCTCTCAAA 672
Db |||
Qy 1078 GATACTAAGCTATTGAAACACACTAGCTATCGGTGACACCATCATCTCAAGAAATTACTA 1137
Db |||
Qy 673 GATACTAAGCTATTGAAACACACTAGCTATCGGTGACACCATCATCTCAAGAAATTACTA 732
Db |||
Qy 1138 GCTCAAGCAAAAGCAATTTAAACAAACCCAGGCTATACGATTTATGAACGTGAC 1197
Db |||
Qy 733 GCTCAAGCAAAAGCAATTTAAACAAACCCAGGCTATACGATTTATGAACGTGAC 792
Db |||
Qy 1198 TCCTCAATCGTCACTCATGACATGACATTTTCGGTACGATTTTACCAATGGATCAAGAG 1257
Db |||
Qy 793 TCCTCAATCGTCACTCATGACATGACATTTTCGGTACGATTTTACCAATGGATCAAGAG 852
Db |||
Qy 1258 TTTACTTACCGTGTAAAAATCGGGAAACAGCTTTATAGGATCAATAAAAAATCTGGTCTG 1317
Db |||
Qy 853 TTTACTTACCGTGTAAAAATCGGGAAACAGCTTTATAGGATCAATAAAAAATCTGGTCTG 912
Db |||
Qy 1318 AATGAAGAAATAAACAACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1377
Db |||
Qy 913 AATGAAGAAATAAACAACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 972
Db |||
Qy 1378 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACTGTTCCACATCAATACGTT 1437
Db |||
Qy 973 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACTGTTCCACATCAATACGTT 1032
Db |||
Qy 1438 GATGTCGATACCAACGAATTTGCTTAAAAAGGTGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 1497
Db |||
Qy 1033 GATGTCGATACCAACGAATTTGCTTAAAAAGGTGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 1092
Db |||

RESULT 4
US-09-940-235-1
; Sequence 1, Application US/09940235
; Publication No. US20030059921A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammara
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/09/940,235
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1245
; TYPE: DNA
; ORGANISM: Streptococcus equisimilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1242)
US-09-940-235-1

Query Match      69.8%; Score 1243.4; DB 3; Length 1245;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1244; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 538 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAAATGGTTGTT 597
Db 1 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAAATGGTTGTT 60
Qy 598 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAATC 657
Db |||
Qy 61 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAATC 120
Db |||
Qy 658 GATCTAAACATCAGACCTGTCTCATGAGGAGAAAGACAGACGAGGCTTAAAGTCCCAATCA 717
Db |||
Qy 121 GATCTAAACATCAGACCTGTCTCATGAGGAGAAAGACAGACGAGGCTTAAAGTCCCAATCA 180
Db |||
Qy 718 AAACATTTGCTACTGATAGTGGCGGATGTCAATAAATCTTGAGAAAGCTGACTACTA 777
Db |||

```

```

Db 181 AAACCAATTTGCTACTGATAGTGGCGGATGTACATAACTTGAGAAAGCTGACTTACTA 240
Qy 778 AAGGCTATTCAAGAAACAATTGATCGCTAAAGTCCACAGTAAACGACGACTACTTTGAGTTC 837
Db 241 AAGGCTATTCAAGAAACAATTGATCGCTAAAGTCCACAGTAAACGACGACTACTTTGAGTTC 300
Qy 838 ATTGATTTTGCAGCGATGCAACCAATTACTGATCGAAACGCGAAGGCTTACTTTGCTGAC 897
Db 301 ATTGATTTTGCAGCGATGCAACCAATTACTGATCGAAACGCGAAGGCTTACTTTGCTGAC 360
Qy 898 AAAGATGGTTTCGGTAACTTCGCGACCAACCTGTCCAAGAAATTTTGTCAAGCGGACAT 957
Db 361 AAAGATGGTTTCGGTAACTTCGCGACCAACCTGTCCAAGAAATTTTGTCAAGCGGACAT 420
Qy 958 GTGCGGTTAGACCATATAAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 1017
Db 421 GTGCGGTTAGACCATATAAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 480
Qy 1018 GAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTCAAA 1077
Db 481 GAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTCAAA 540
Qy 1078 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTA 1137
Db 541 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTA 600
Qy 1138 GCTCAAGCAAAAGCAATTTTAAACAAAACCAACCCAGGCTATACGATTTATGAACGTGAC 1197
Db 601 GCTCAAGCAAAAGCAATTTTAAACAAAACCAACCCAGGCTATACGATTTATGAACGTGAC 660
Qy 1198 TCCTCAATCGTACTCATGACAAATGATTTTCGTTACGATTTTACCAATGATCAAGAG 1257
Db 661 TCCTCAATCGTACTCATGACAAATGATTTTCGTTACGATTTTACCAATGATCAAGAG 720
Qy 1258 TTCTACTTACCGTGTAAATTCGGACACAGCTTATAGGATCAATAAATCTGGTCTG 1317
Db 721 TTCTACTTACCGTGTAAATTCGGACACAGCTTATAGGATCAATAAATCTGGTCTG 780
Qy 1318 AATGAAGAAATAAACCACTGACCTGATCTCTGAGAAATATTAAGTCTTAAAAAGGG 1377
Db 781 AATGAAGAAATAAACCACTGACCTGATCTCTGAGAAATATTAAGTCTTAAAAAGGG 840
Qy 1378 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTTGAAACTGTTTACCACCAATACGTT 1437
Db 841 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTTGAAACTGTTTACCACCAATACGTT 900
Qy 1438 GATGTCGATACCAACGAATGCTTAAAGAGTGGAGAGCTCTTAAAGCTAGGAAACGTAAC 1497
Db 901 GATGTCGATACCAACGAATGCTTAAAGAGTGGAGAGCTCTTAAAGCTAGGAAACGTAAC 960
Qy 1498 TTAGACTTCAGAGATTTATACGATCTCTGATAAGGCTAAACTTACTCTACAACTCTC 1557
Db 961 TTAGACTTCAGAGATTTATACGATCTCTGATAAGGCTAAACTTACTCTACAACTCTC 1020
Qy 1558 GATGCTTTTGGTATTATGACTATACCTTAACTTGAAGAGTGAAGGATAATCAGCATGAC 1617
Db 1021 GATGCTTTTGGTATTATGACTATACCTTAACTTGAAGAGTGAAGGATAATCAGCATGAC 1080
Qy 1618 ACCAAGCGTATCATACCGTTTATATGGCAAGCGACCCGAAGAGAGAGATGCTAGCTAT 1677
Db 1081 ACCAAGCGTATCATACCGTTTATATGGCAAGCGACCCGAAGAGAGAGATGCTAGCTAT 1140
Qy 1678 CATTTAGCTTATGATAAGATCGTTTATACCGAAGAGAGAAAGTTTACAGTACCTG 1737
Db 1141 CATTTAGCTTATGATAAGATCGTTTATACCGAAGAGAGAAAGTTTACAGTACCTG 1200
Qy 1738 CGTTATACAGGAGACCTTATACCTGATTAACCCCTTAAACGACAAATAA 1782
Db 1201 CGTTATACAGGAGACCTTATACCTGATTAACCCCTTAAACGACAAATAA 1245

```

RESULT 5

US-09-940-235-6

```

; Sequence 6, Application US/09940235
; Publication No. US20030059921A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kamara
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/09/940,235
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1327
; TYPE: DNA
; ORGANISM: Streptococcus equisimilis
US-09-940-235-6

```

```

Query Match 69.1%; Score 1232.2; DB 3; Length 1327;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1237; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```

```

Qy 538 ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCACAAACGAGCAATGTTGTT 597
Db 83 ATAGCTGCTCTGAATGGCTACTAGATCGTCTCTGTAAATAACAGCCNAATGTTGTT 142
Qy 598 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTCAAATC 657
Db 143 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTCAAATC 202
Qy 658 GATCTAATCATCAACCTGCTCATGAGGAAAGACAGACGAGGCTTAAGTCCAAATCA 717
Db 203 GATCTAATCATCAACCTGCTCATGAGGAAAGACAGACGAGGCTTAAGTCCAAATCA 262
Qy 718 AAACCATTTGCTACTGATAGTGGCGGATGTCAATAAACTTGAGAAAGCTGACTTACTA 777
Db 263 AAACCATTTGCTACTGATAGTGGCGGATGTCAATAAACTTGAGAAAGCTGACTTACTA 322
Qy 778 AAGGCTATTCAAGAAACAATTCATCGTAAACGTCACAGTAAACGACGACTCTTTGAGGTC 837
Db 323 AAGGCTATTCAAGAAACAATTCATCGTAAACGTCACAGTAAACGACGACTCTTTGAGGTC 382
Qy 838 ATTGATTTTGAACGCGATGCAACCATTTACTGATGAAACGCGAAGGCTACTTTGCTGAC 897
Db 383 ATTGATTTTGAACGCGATGCAACCATTTACTGATGAAACGCGAAGGCTACTTTGCTGAC 442
Qy 898 AAAGATGGTTTCGGTAACTTCGCGACCCAACTGTCGAAAGATTTTCTAAGCGGACAT 957
Db 443 AAAGATGGTTTCGGTAACTTCGCGACCCAACTGTCGAAAGATTTTCTAAGCGGACAT 502
Qy 958 GTGCGGTTAGACCATATAAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 1017
Db 503 GTGCGGTTAGACCATATAAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 562
Qy 1018 GAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTCAAA 1077
Db 563 GAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTCAAA 622
Qy 1078 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTA 1137
Db 623 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTA 682

```



```

QY 1138 GCTCAAGCAAAAGCAATTTTAAACAAACAAACCCAGGCTATAGATTTATGAACGTGAC 1197
Db |||||||
QY 1198 TCTCAATCGTCACTCATGCAATGATGATTTTCGGTACGATTTTACCAATGGATCAAGAG 1257
Db |||||||
QY 1258 TTTACTTACCGTGTAAATAATCGGAAACAAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 1317
Db |||||||
QY 1318 AATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 1377
Db |||||||
QY 1378 GAAAGCCGTATGATCCCTTTGATCGAGTCACTTGAAACTGTTTCAACATCAAAATAGCTT 1437
Db |||||||
QY 1438 GATGCGATACCAACCAATGCTTAAAGGTGACGAGCTCTTAAAGCTAGCGAACGTAAC 1497
Db |||||||
QY 1498 TTAGACTTCAGAGATTTATACGATCTCGTGATAAGGCTTAACTCTCTCAACAATCTC 1557
Db |||||||
QY 1558 GATGCTTTGGTATTATGACATATACCTTAACTGGAAAAAGTAGAGGATATACGATGAC 1617
Db |||||||
QY 1618 ACCAACCGTATCATACCTTTATATGGGCAAGCGACCGAAGGAGAGATGCTAGCTAT 1677
Db |||||||
QY 1678 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAGACGAGAAAGTTTACAGCTACCTG 1737
Db |||||||
QY 1738 CTTTATACAGGACACCTATACCTGATAACCCCTTAACGACAAATATA 1782
Db |||||||
QY 1782 CTTTATACAGGACACCTATACCTGATAACCCCTTAACGACAAATATA 1822
Db |||||||

```

RESULT 6

```

US-09-940-235-10
; Sequence 10, Application US/09940235
; Publication No. US20030059921A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kamnara
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/09/940,235
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 1661
; TYPE: DNA

```

```

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hybrid cassette
US-09-940-235-10

```

```

Query Match 64.5%; Score 1150.2; DB 3; Length 1661;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1152; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 538 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAACAGCAATTTGGTT 597
Db |||||||
QY 598 AGCGTTCCTGCTACTGTTGAGGGGCAATCAAGACATTAAGTCTTAATTTTGAATC 657
Db |||||||
QY 658 GATCTAACATCAACGACCTGCTCATGGAGGAAAGACAGACGAGCAAGGCTTTAAGTCCAAATCA 717
Db |||||||
QY 718 AACCACTTTGCTACTGATGAGCGGATGTCAATAACTTTGAGAAAGCTGACTTACTA 777
Db |||||||
QY 778 AAGGCTATTCAAGNACAAATTTGATCGCTAAAGCTCCACAGTAACGACGACTACTTTGAGT 837
Db |||||||
QY 838 ATTGATTTTCAAGAGCGATGCAACCATTTACTGATCGAAACGCAAGGCTTACTTTGCTGAC 897
Db |||||||
QY 898 AAGAGTGGTTCCGTTAACTTTGCCGACCCCAACCTGTCCAGAAATTTTGTCTAAGCGGACAT 957
Db |||||||
QY 958 GTCCGGTTAGACCATATAAAGAAACCAATATACAAACCAAGCGAAATCTGTTGATGTG 1017
Db |||||||
QY 1018 GAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTACAGCAGGCTCTCAA 1077
Db |||||||
QY 1078 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTAATA 1137
Db |||||||
QY 1138 GCTCAAGCAAAAGCAATTTTAAACAAACCAACCCAGGCTATAGATTTATGAACGTCGAC 1197
Db |||||||
QY 1198 TCCTCAATCGTCACTCATGACATGATTTTCCGTACGATTTTACCAATGGATCAAGAG 1257
Db |||||||
QY 1258 TTTTACTTACCGTGTAAAAATCGGAAACAAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 1317
Db |||||||
QY 1318 AATGAAGAAATAAAACAACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 1377
Db |||||||
QY 1378 GAAAGCCGTATGATCCCTTTGATCGAGTCACTTGAACCTGTTTCAACATCAAAATAGCTT 1437
Db |||||||
QY 1438 GATGTCGATACCAACCAATGCTTAAAGGTGACGAGCTCTTAAAGCTAGCGAACGTAAC 1497
Db |||||||
QY 1498 GATGTCGATACCAACCAATGCTTAAAGGTGACGAGCTCTTAAAGCTAGCGAACGTAAC 1542
Db |||||||
QY 1542 GATGTCGATACCAACCAATGCTTAAAGGTGACGAGCTCTTAAAGCTAGCGAACGTAAC 1602
Db |||||||
QY 1602 GATGTCGATACCAACCAATGCTTAAAGGTGACGAGCTCTTAAAGCTAGCGAACGTAAC 1662
Db |||||||
QY 1662 GATGTCGATACCAACCAATGCTTAAAGGTGACGAGCTCTTAAAGCTAGCGAACGTAAC 1722
Db |||||||
QY 1722 GATGTCGATACCAACCAATGCTTAAAGGTGACGAGCTCTTAAAGCTAGCGAACGTAAC 1782
Db |||||||
QY 1782 GATGTCGATACCAACCAATGCTTAAAGGTGACGAGCTCTTAAAGCTAGCGAACGTAAC 1842
Db |||||||
QY 1842 GATGTCGATACCAACCAATGCTTAAAGGTGACGAGCTCTTAAAGCTAGCGAACGTAAC 1902
Db |||||||
QY 1902 GATGTCGATACCAACCAATGCTTAAAGGTGACGAGCTCTTAAAGCTAGCGAACGTAAC 1962
Db |||||||
QY 1962 GATGTCGATACCAACCAATGCTTAAAGGTGACGAGCTCTTAAAGCTAGCGAACGTAAC 2022
Db |||||||
QY 2022 GATGTCGATACCAACCAATGCTTAAAGGTGACGAGCTCTTAAAGCTAGCGAACGTAAC 2082
Db |||||||
QY 2082 GATGTCGATACCAACCAATGCTTAAAGGTGACGAGCTCTTAAAGCTAGCGAACGTAAC 2142
Db |||||||

```


QY 1498 TTAGACTTCAGAGATTATACGATCCTCGTGTAAGGCTAAACCTACTCTCAACAATCTC 1557
DB 1143 TTAGACTTCAGAGATTATACGATCCTCGTGTAAGGCTAAACCTACTCTCAACAATCTC 1202
QY 1558 GATGCTTTTGGTATTATGACTACTACCTTAACCTGGAAGCTAGAGGATATACGATGAC 1617
DB 1203 GATGCTTTTGGTATTATGACTACTACCTTAACCTGGAAGCTAGAGGATATACGATGAC 1262
QY 1618 ACCAACCGTATCATACACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAAATGCTAGCTAT 1677
DB 1263 ACCAACCGTATCATACACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAAATGCTAGCTAT 1322
QY 1678 CATTTAGCCGTATGAT 1692
DB 1323 CATTTAGCCGTGGT 1337

RESULT 7

US-09-940-235-9
; Sequence 9, Application US/09940235
; Publication No. US20030059921A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammarra
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/09/940, 235
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1541
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hybrid cassette
US-09-940-235-9

Query Match 1 63.6%; Score 1134.2; DB 3; Length 1541;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1142; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 538 ATTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACGACCAATGGTGT 597
DB 33 ATAGCTGGTCCGTAATGGCTACTAGATCGCTCTCTGTAATAACAGCAATGGTGT 92
QY 598 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATAGTCTTAAATTTTGAATC 657
DB 93 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATAGTCTTAAATTTTGAATC 152
QY 658 GATCTAAATCATACGACCTGCTCATGGAGGAAGACAGACAGGCTTAAAGTCCAAATCA 717
DB 153 GATCTAAATCATACGACCTGCTCATGGAGGAAGACAGACAGGCTTAAAGTCCAAATCA 212
QY 718 AAACCATTTGCTACTGATAGTGGCGGATGTCAATAAATTCGAGAAAGCTGACTACTA 777
DB 213 AAACCATTTGCTACTGATAGTGGCGGATGTCAATAAATTCGAGAAAGCTGACTACTA 272
QY 778 AAGGCTATTCAAGAACAAATTTGATCGCTTAACGCTCCACAGTAACGACACTACTTTGAGGTC 837
DB 273 AAGGCTATTCAAGAACAAATTTGATCGCTTAACGCTCCACAGTAACGACACTACTTTGAGGTC 332

QY 838 ATTGATTTTGAAGCGATGCAACCAATTACTGATCGAAACGGCAAGGCTCTACTTTGCTGAC 897
DB 333 ATTGATTTTGAAGCGATGCAACCAATTACTGATCGAAACGGCAAGGCTCTACTTTGCTGAC 392
QY 898 AAAGATGGTTGCGTAAACCTTGCCTGCAACCAACCTGTCGCAAGAAATTTTTCCTAAGCGGACAT 957
DB 393 AAAGATGGTTGCGTAAACCTTGCCTGCAACCAACCTGTCGCAAGAAATTTTTCCTAAGCGGACAT 452
QY 958 GTGCGCGTTAGACCATATAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 1017
DB 453 GTGCGCGTTAGACCATATAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 512
QY 1018 GAATATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTCAGACCAAGTCTCAAA 1077
DB 513 GAATATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTCAGACCAAGTCTCAAA 572
QY 1078 GATCTAAGCTATTGAAACCACTAGCTATCGGTGACACCATCACATCTCAAGAAATTA 1137
DB 573 GATCTAAGCTATTGAAACCACTAGCTATCGGTGACACCATCACATCTCAAGAAATTA 632
QY 1138 GCTCAAGCACAAGCATTTTAAACCAACCAACCAAGCGCTATACGATTTATGAACGTGAC 1197
DB 633 GCTCAAGCACAAGCATTTTAAACCAACCAACCAAGCGCTATACGATTTATGAACGTGAC 692
QY 1198 TCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATCGATCAAGAG 1257
DB 693 TCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATCGATCAAGAG 752
QY 1258 TTTACTTACCGTGTAAAAATTCGGGAAACCAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 1317
DB 753 TTTACTTACCGTGTAAAAATTCGGGAAACCAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 812
QY 1318 AATGAAGAAATAAACAACACCTGATCTCTGAGAAATATATACGCTTTAAAAAAGGG 1377
DB 813 AATGAAGAAATAAACAACACCTGATCTCTGAGAAATATATACGCTTTAAAAAAGGG 872
QY 1378 GAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCCACCATCAATACGTT 1437
DB 873 GAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCCACCATCAATACGTT 932
QY 1438 GATGTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAAGCTAAC 1497
DB 933 GATGTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAAGCTAAC 992
QY 1498 TTAGACTTCAGAGATTTTATACGATCTCTGTAAGGCTAAACCTACTCTCAACAATCTC 1557
DB 993 TTAGACTTCAGAGATTTTATACGATCTCTGTAAGGCTAAACCTACTCTCAACAATCTC 1052
QY 1558 GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAGAGTAGAGGATAATACAGATGAC 1617
DB 1053 GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAGAGTAGAGGATAATACAGATGAC 1112
QY 1618 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAAATGCTAGCTAT 1677
DB 1113 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAAATGCTAGCTAT 1172
QY 1678 CATTTAGCCGTATGAT 1692
DB 1173 CATTTAGCCGTGGT 1187

RESULT 8

US-10-474-792-657
; Sequence 657, Application US/10474792
; Publication No. US20040236072A1
; GENERAL INFORMATION:
; APPLICANT: Olmsted, Stephen
; APPLICANT: Zagursky, Robert
; APPLICANT: Nickbarg, Elliot
; APPLICANT: Winter, Lourie
; TITLE OF INVENTION: SURFACE PROTEINS OF STREPTOCOCCUS PYOGENES
; FILE REFERENCE: AM 100399

; CURRENT APPLICATION NUMBER: US/10/474,792
 ; CURRENT FILING DATE: 2003-10-14
 ; NUMBER OF SEQ ID NOS: 674
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 657
 ; LENGTH: 1323
 ; TYPE: DNA
 ; ORGANISM: Streptococcus pyogenes
 US-10-474-792-657

Query Match 60.7%; Score 1081.2; DB 8; Length 1323;
 Best Local Similarity 91.7%; Pred. No. 1.3e-299;
 Matches 1143; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

Qy	537	TATTGCTGGACCTGAGTGGCTGTAGACGGCTCATCTGTCAACAAACAGCGCAATGGTTGT	596
Db	78	TATTGCTGGTATGATGGCTACAGACCGTCCACCTATCAATCAACAGCGGTTAGTTGT	137
Qy	597	TAGCGTTGCTGTTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTGAAAT	656
Db	138	TAGTATGGCCGGTATCGTTGAAGGTACCGATAAAAAGTTTTTATAAATTTTTTTGAAAT	197
Qy	657	CGATCTAACATCAACACCTGCTCAGGAGGAAAGACAGAGCAAGGCTTAAGTCCAAAATC	716
Db	198	CGATCTAACATCAACACCTGCTCAGGAGGAAAGACAGAGCAAGGCTTAAGTCCAAAATC	257
Qy	717	AAACCAATTTGCTACTGATAGTGGCGGATGTACACATAAATTTGAGAAAGCTGACTTACT	776
Db	258	AAACCAATTTGCTACAGATAATGGCGCAATGCCACATAAATTTGAAAGAGCTGACTTATT	317
Qy	777	AAAGCTATTCAAGACCAATTTGATCGCTTAACGCTCCACAGTAACAGCGACTTACTTTGAGGT	836
Db	318	AAAGCTATTCAAGACCAATTTGATCGCTTAACGCTCCACAGTAACAGCGACTTACTTTGAGGT	377
Qy	837	CATTGATTTGCAAGCAATGCAACCAATTAAGTCAAGAAACCGCAAGGTCTACTTTGCTGA	896
Db	378	CATTGATTTGCAAGCAATGCAACCAATTAAGTCAAGAAACCGCAAGGTCTACTTTGCTGA	437
Qy	897	CAAGATGGTTCGGTAACTTCCCGACCCCAACCTGTCCAGAAATTTTGTGAAGCGGACA	956
Db	438	CAAGATGGTTCGGTAACTTCCCGACCCCAACCTGTCCAGAAATTTTGTGAAGCGGACA	497
Qy	957	TGTGCGGCTTAGACCATATAAAGAAAAACCAATACAAAAACCAAGCGAAATCTCTGATGT	1016
Db	498	TGTGCGGCTTAGACCATATAAAGAAAAACCAATACAAAAACCAAGCGAAATCTCTGATGT	557
Qy	1017	GGAATATATGTPACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCGGCTCAA	1076
Db	558	GGAATATATGTPACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCGGCTCAA	617
Qy	1077	AGATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATTACT	1136
Db	618	AGATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATTACT	677
Qy	1137	AGCTCAAGCACAAAGCATTTTAAACAAAAACCAACCGGCTATACGATTTATCAACGTTGA	1196
Db	678	AGCTCAAGCACAAAGCATTTTAAACAAAAACCAACCGGCTATACGATTTATCAACGTTGA	737
Qy	1197	CTCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGATCAAGA	1256
Db	738	CTCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGATCAAGA	797
Qy	1257	GTTTACTTACCGTGTAAATAATCGGGAAACAGCTTATAGATCAATAAATAATCTCGGTCT	1316
Db	798	GTTTACTTACCGTGTAAATAATCGGGAAACAGCTTATAGATCAATAAATAATCTCGGTCT	857
Qy	1317	GAACTAAGAAAAATAACACCTGATCTCTGAGAAATATATAGTCTCTTAAATAAGG	1376
Db	858	TAAAGAAAAACGAAACACCTGATCTGGTCTCTGAGAAATATATAGTCTCTTAAATAAGG	917
Qy	1377	GGAAAGCGGTATGATCCCTTTTGTATCGAGTCACTTTGAAACTGTTCCACCATCAATACGT	1436
Db	918	GGAAAGCGGTATGATCCCTTTTGTATCGAGTCACTTTGAAACTGTTCCACCATCAATACGT	977

RESULT 9

US-10-210-120-49
 ; Sequence 49, Application US/10210120
 ; Publication No. US20030175736A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Chinnaiyan, Arul M.
 ; APPLICANT: Rubin, Mark A.
 ; APPLICANT: Sreekumar, Arun
 ; TITLE OF INVENTION: Expression Profile of Prostate Cancer
 ; FILE REFERENCE: UM-07221
 ; CURRENT APPLICATION NUMBER: US/10/210,120
 ; CURRENT FILING DATE: 2002-08-01
 ; PRIOR APPLICATION NUMBER: US 60/309,581
 ; PRIOR FILING DATE: 2001-08-02
 ; PRIOR APPLICATION NUMBER: US 60/334,468
 ; PRIOR FILING DATE: 2001-11-15
 ; NUMBER OF SEQ ID NOS: 123
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 49
 ; LENGTH: 2127
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-210-120-49

Query Match 13.1%; Score 232.8; DB 6; Length 2127;
 Best Local Similarity 79.8%; Pred. No. 2e-55;
 Matches 308; Conservative 0; Mismatches 32; Indels 46; Gaps 1;

Qy	254	CCCATAGCTCAGAAAGTCTTTTGTATCATGCTGGGACTTCTCTATGTGGTGGGAAACG	313
Db	597	CCCATAGCTCAGAAAGTCTTTTGTATCATGCTGGGACTTCTCTATGTGGTGGGAAACG	656
Qy	314	-----TGGGAGAAAGGCAGC	327
Db	657	TGGGAGAAAGCCCTACCAAGGCTGGATGATGTTGCTTGGGAGAGGCGAGC	716
Qy	328	GGACCGCATCTTGGCACTTCTAGAAATAGATGCAACGATCAGGACACAAAGGACATCTTAT	387
Db	717	GGACCGCATCTTGGCACTTCTAGAAATAGATGCAACGATCAGGACACAAAGGACATCTTAT	776
Qy	388	AGATTTGGACACCTGGGACCAAGAGTAATCGAGAAACCTGCTCCAGTGCATCTGC	447
Db	777	AGATTTGGACACCTGGGACCAAGAGTAATCGAGAAACCTGCTCCAGTGCATCTGC	836
Qy	448	ACAGGCAACCGGCGAGGAGAGTGAAGTGTGAGAGGCACACCTCTGTGAGACCAACATCG	507
Db	837	ACAGGCAACCGGCGAGGAGAGTGAAGTGTGAGAGGCACACCTCTGTGAGACCAACATCG	896

QY 508 AGCGGATCTGGCCCTTACCGATGTTCTGATTTCTGAGCCTGAGTGGCTGCTGACACGT 567
 Db |||||
 QY 897 AGCGGATCTGGCCCTTACCGATGTTCTGATTTCTGAGCCTGAGTGGCTGCTGACACGT 956
 Db |||||
 QY 568 CCATCTGTCAACAACAGCCCAATTTGGT 593
 Db |||||
 QY 957 CAGCCTCTCCCTATGCGCCACTGTGT 982
 Db |||||

RESULT 10

US-10-956-157-4288
 ; Sequence 4288, Application US/10956157
 ; Publication No. US20050118625A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wyeth
 ; APPLICANT: Mounts, William
 ; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
 ; FILE OF INVENTION: HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
 ; CURRENT FILING DATE: 2004-10-04
 ; NUMBER OF SEQ ID NOS: 319805
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 4288
 ; LENGTH: 2127
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-956-157-4288

Query Match 13.1%; Score 232.8; DB 9; Length 2127;
 Best Local Similarity 79.8%; Pred. No. 2e-55;
 Matches 308; Conservative 0; Mismatches 32; Indels 46; Gaps 1;

QY 254 CCCATAGCTGAGAGTGTGTTTGTATCATGCTGCTGGGACTTCTTATGTTGGTGGGAGAAACG 313
 Db |||||
 QY 597 CCCATAGCTGAGAGTGTGTTTGTATCATGCTGCTGGGACTTCTTATGTTGGTGGGAGAAACG 656
 Db |||||
 QY 314 -----TGGGAGAGGCGAGC 327
 Db |||||
 QY 657 TGGGAGAGGCGCTACCAAGCTGGATGATGTTGTTGCTTGGGAGAGGCGAGC 716
 Db |||||
 QY 328 GGACGCATCATCTTGCATCTTCTAGAAATAGATGCAACGATCAGGACACAAAGGACATCTTAT 387
 Db |||||
 QY 717 GGACGCATCATCTTGCATCTTCTAGAAATAGATGCAACGATCAGGACACAAAGGACATCTTAT 776
 Db |||||
 QY 388 AGAATTGGAGACACCTGGAGCAAGAGGTAATCGAGAAACCTGCTCCAGTGCATCTGC 447
 Db |||||
 QY 777 AGAATTGGAGACACCTGGAGCAAGAGGTAATCGAGAAACCTGCTCCAGTGCATCTGC 836
 Db |||||
 QY 448 ACAGGCAACGCGGAGGAGTGAAGTGTGAGAGGACACCTCTGTGAGACCAACATCG 507
 Db |||||
 QY 837 ACAGGCAACGCGGAGGAGTGAAGTGTGAGAGGACACCTCTGTGAGACCAACATCG 896
 Db |||||
 QY 508 AGCGGATCTGGCCCTTACCGATGTTCTGATTTCTGAGCCTGAGTGGCTGCTGACACGT 567
 Db |||||
 QY 897 AGCGGATCTGGCCCTTACCGATGTTCTGATTTCTGAGCCTGAGTGGCTGCTGACACGT 956
 Db |||||
 QY 568 CCATCTGTCAACAACAGCCCAATTTGGT 593
 Db |||||
 QY 957 CAGCCTCTCCCTATGCGCCACTGTGT 982
 Db |||||

RESULT 11

US-10-909-035-49
 ; Sequence 49, Application US/10909035
 ; Publication No. US20050136493A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rubin, Mark A.
 ; APPLICANT: Chinnaiyan, Arul M.
 ; APPLICANT: Laxman, Bharathi
 ; APPLICANT: Sreekumar, Arun
 ; TITLE OF INVENTION: AMACR Cancer Markers

FILE REFERENCE: UM-09098
 ; CURRENT APPLICATION NUMBER: US/10/909,035
 ; CURRENT FILING DATE: 2004-07-30
 ; NUMBER OF SEQ ID NOS: 123
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 49
 ; LENGTH: 2127
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-909-035-49

Query Match 13.1%; Score 232.8; DB 9; Length 2127;
 Best Local Similarity 79.8%; Pred. No. 2e-55;
 Matches 308; Conservative 0; Mismatches 32; Indels 46; Gaps 1;

QY 254 CCCATAGCTGAGAGTGTGTTTGTATCATGCTGCTGGGACTTCTTATGTTGGTGGGAGAAACG 313
 Db |||||
 QY 597 CCCATAGCTGAGAGTGTGTTTGTATCATGCTGCTGGGACTTCTTATGTTGGTGGGAGAAACG 656
 Db |||||
 QY 314 -----TGGGAGAGGCGAGC 327
 Db |||||
 QY 657 TGGGAGAGGCGCTACCAAGCTGGATGATGTTGTTGCTTGGGAGAGGCGAGC 716
 Db |||||
 QY 328 GGACGCATCATCTTGCATCTTCTAGAAATAGATGCAACGATCAGGACACAAAGGACATCTTAT 387
 Db |||||
 QY 717 GGACGCATCATCTTGCATCTTCTAGAAATAGATGCAACGATCAGGACACAAAGGACATCTTAT 776
 Db |||||
 QY 388 AGAATTGGAGACACCTGGAGCAAGAGGTAATCGAGAAACCTGCTCCAGTGCATCTGC 447
 Db |||||
 QY 777 AGAATTGGAGACACCTGGAGCAAGAGGTAATCGAGAAACCTGCTCCAGTGCATCTGC 836
 Db |||||
 QY 448 ACAGGCAACGCGGAGGAGTGAAGTGTGAGAGGACACCTCTGTGAGACCAACATCG 507
 Db |||||
 QY 837 ACAGGCAACGCGGAGGAGTGAAGTGTGAGAGGACACCTCTGTGAGACCAACATCG 896
 Db |||||
 QY 508 AGCGGATCTGGCCCTTACCGATGTTCTGATTTCTGAGCCTGAGTGGCTGCTGACACGT 567
 Db |||||
 QY 897 AGCGGATCTGGCCCTTACCGATGTTCTGATTTCTGAGCCTGAGTGGCTGCTGACACGT 956
 Db |||||
 QY 568 CCATCTGTCAACAACAGCCCAATTTGGT 593
 Db |||||
 QY 957 CAGCCTCTCCCTATGCGCCACTGTGT 982
 Db |||||

RESULT 12

US-10-741-601-70
 ; Sequence 70, Application US/10741601
 ; Publication No. US20040166519A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele et al.
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; FILE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001500
 ; CURRENT APPLICATION NUMBER: US/10/741,601
 ; CURRENT FILING DATE: 2003-12-22
 ; NUMBER OF SEQ ID NOS: 26415
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 70
 ; LENGTH: 2443
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-741-601-70

Query Match 13.1%; Score 232.8; DB 7; Length 2443;
 Best Local Similarity 79.8%; Pred. No. 2.1e-55;
 Matches 308; Conservative 0; Mismatches 33; Indels 46; Gaps 1;

QY 254 CCCATAGCTGAGAGTGTGTTTGTATCATGCTGCTGGGACTTCTTATGTTGGTGGGAGAAACG 313
 Db |||||
 QY 913 CCCATAGCTGAGAGTGTGTTTGTATCATGCTGCTGGGACTTCTTATGTTGGTGGGAGAAACG 972
 Db |||||
 QY 314 -----TGGGAGAGGCGAGC 327
 Db |||||

Query Match	13.1%;	Score	232.8;	DB	8;	Length	2488;	
Best Local Similarity	79.8%;	Pred.	No. 2.1e-55;					
Matches	308;	Conservative	0;	Mismatches	32;	Indels	46; Gaps	1;

QY	254	CCCATAGCTCAGAGAAGTGTGTTTGATCATGCTGCTGGGACTTCCTATGTGTGGTCGGAGAAACG	313
Dd	913	CCCNAGCTCGAAGAAGTGTGTTTGATCATGCTGCTGGGACTTCCTATGTGTGTGGAGAAACG	972
QY	314	-----TGGGAGAAGGCAGC	327
Dd	973	TGGGAGAAGGCCATTCAAGGCTGGATGATGGTAGATTGTACTTTCCTGGGAGAAGGCAGC	1032
QY	328	GGAGCGCATCTTGGCATCTTCTAGAAATAGATGCAACGATCAGACACACAAGGACATCCTAT	387
Dd	1033	GGAAGCATCTTGGCACTTCTAGAAATAGATGCAACGATCAGGACACAAGGACATCCTAT	1092
QY	388	AGAAATTGGAGACACCTGGAGCAAGAAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC	447
Dd	1093	AGAAATTGGAGACACCTGGAGCAAGAAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC	1152
QY	448	ACAGGCNAACGGCCGAGGAGAGTGGAGTGTGAGAGGCACACCTCTGTGCAGACCACATCG	507
Dd	1153	ACAGGCNAACGGCCGAGGAGAGTGGAAAGTGTGAGAGGCACACCTCTGTGCAGACCACATCG	1212
QY	508	AGCGGATCTGGCCCCCTTACCGATGTTTCGGTATTGCTGCAGCTCAGTGGCTGTCTAGACCGT	567
Dd	1213	AGCGGATCTGGCCCCCTTACCGATGTTTCGGTGCAGCTGTTTACCNAACCGAGCCTCACCCC	1272
QY	568	CNATCTGTCAACAAACAGCCAAATGGT	593
Dd	1273	CAGCCTCTCTCCCTATGGCCACTGTGT	1298

Search completed: February 1, 2006, 14:21:21
Job time : 1474.42 secs

Result No.	Score			Query		DB	ID	Description
	Score	Match	Length	Match	Length			
1	1782	100.0	1782	7	US-10-631-558-11	Sequence 11, Appl		
2	1684	94.5	2096	7	US-10-631-558-12	Sequence 12, Appl		
3	1245	69.9	1377	7	US-10-631-558-5	Sequence 5, Appl		
4	1243.4	69.8	1245	7	US-10-631-558-1	Sequence 1, Appl		
5	1232.2	69.1	1327	7	US-10-631-558-6	Sequence 6, Appl		
6	1150.2	64.5	1661	7	US-10-631-558-10	Sequence 10, Appl		
7	1134.2	63.6	1541	7	US-10-631-558-9	Sequence 9, Appl		
8	232.8	13.1	2443	7	US-10-995-561-114	Sequence 114, Appl		
9	232.8	13.1	2488	7	US-10-995-561-105	Sequence 105, Appl		
10	232.8	13.1	6510	7	US-10-995-561-112	Sequence 112, Appl		
11	232.8	13.1	7823	7	US-10-995-561-117	Sequence 117, Appl		
12	232.8	13.1	7848	7	US-10-995-561-111	Sequence 111, Appl		
13	232.8	13.1	7935	7	US-10-995-561-113	Sequence 113, Appl		
14	232.8	13.1	7959	7	US-10-995-561-108	Sequence 108, Appl		
15	232.8	13.1	8013	7	US-10-995-561-104	Sequence 104, Appl		
16	232.8	13.1	8155	7	US-10-995-561-116	Sequence 116, Appl		
17	232.8	13.1	8266	7	US-10-995-561-107	Sequence 107, Appl		
18	232.8	13.1	8232	7	US-10-821-234-693	Sequence 693, Appl		
19	232.8	13.1	8278	7	US-10-995-561-106	Sequence 106, Appl		
20	232.8	13.1	8332	7	US-10-995-561-110	Sequence 110, Appl		
21	232.8	13.1	8371	7	US-10-995-561-109	Sequence 109, Appl		
22	228	12.8	7777	7	US-10-631-558-3	Sequence 3, Appl		

Db 61 CAGCCTAGCGGGTCTCAACGACAGGACGATCATGCGCACCGGTGCGGACCCCA 120
Qy 121 AGCGTCCCGAGATCTCGATCCGCGAAATTAATACGACTCACTATAGGGAGACCAAC 180
Db 121 AGCGTCCCGAGATCTCGATCCGCGAAATTAATACGACTCACTATAGGGAGACCAAC 180
Qy 181 GGTTCCTCTAGAAATAATTTTGTAACTTTAAGAAAGGAGATATACCATGTCGAAGC 240
Db 181 GGTTCCTCTAGAAATAATTTTGTAACTTTAAGAAAGGAGATATACCATGTCGAAGC 240
Qy 241 ACAACAGATTGTACCATAGCTGAGAAAGTGTTCGATCATGCTGGGACTTCTTATGT 300
Db 241 ACAACAGATTGTACCATAGCTGAGAAAGTGTTCGATCATGCTGGGACTTCTTATGT 300
Qy 301 GGTTCGAGAAACGTGGGAGAGCGGAGCGGATCATCTTGCATCTTGAATAATAGATGC 360
Db 301 GGTTCGAGAAACGTGGGAGAGCGGAGCGGATCATCTTGCATCTTGAATAATAGATGC 360
Qy 361 AACGATCAGGACACAAAGGACATCTATAGAAATTTGGAGACACCTGGAGCAAGAGGATAAT 420
Db 361 AACGATCAGGACACAAAGGACATCTATAGAAATTTGGAGACACCTGGAGCAAGAGGATAAT 420
Qy 421 CGAGAAACCTGCTCCAGTGCATCTGCACAGGCAACCGCGGAGGAGTGGAAAGTGTGAG 480
Db 421 CGAGAAACCTGCTCCAGTGCATCTGCACAGGCAACCGCGGAGGAGTGGAAAGTGTGAG 480
Qy 481 AGGCACACCTCTGTGCGAGACCAATCGAGCGGATCTGGCCCTTCCACGATGTTGTAAT 540
Db 481 AGGCACACCTCTGTGCGAGACCAATCGAGCGGATCTGGCCCTTCCACGATGTTGTAAT 540
Qy 541 GCTGACCTGAGTGGCTGTAGACCTTCCATCTGTCAACAGCAAACTTGGTGTGAGC 600
Db 541 GCTGACCTGAGTGGCTGTAGACCTTCCATCTGTCAACAGCAAACTTGGTGTGAGC 600
Qy 601 GTTGCTGTGTAATCTGAGGGGAGCAATCAAGACATTAAGTCTTAAATTTTTTGAATCGAT 660
Db 601 GTTGCTGTGTAATCTGAGGGGAGCAATCAAGACATTAAGTCTTAAATTTTTTGAATCGAT 660
Qy 661 CTAAACATCAGACCTGCTCATGAGGAAAGACAGAGCAAGGCTTAAGTCCAAATCAAAA 720
Db 661 CTAAACATCAGACCTGCTCATGAGGAAAGACAGAGCAAGGCTTAAGTCCAAATCAAAA 720
Qy 721 CCATTTGCTACTGATAGTGGCGGATGTCAATTAATCTGAGAAAGCTACTTAAAG 780
Db 721 CCATTTGCTACTGATAGTGGCGGATGTCAATTAATCTGAGAAAGCTACTTAAAG 780
Qy 781 GCTATTCAAGAAACATTTGATCGTAACTGCTCAAGTAAACGACCACTACTTTGAGGTCAAT 840
Db 781 GCTATTCAAGAAACATTTGATCGTAACTGCTCAAGTAAACGACCACTACTTTGAGGTCAAT 840
Qy 841 GATTTTGAAGGATGCAACCACTACTGATCGAAACGGCAAGTCTACTTTGCTGCAAAA 900
Db 841 GATTTTGAAGGATGCAACCACTACTGATCGAAACGGCAAGTCTACTTTGCTGCAAAA 900
Qy 901 GATGTTTGGTAACTTGGCGACCCCACTGCTCAAGAAATTTTGTAAAGCGGACATGTG 960
Db 901 GATGTTTGGTAACTTGGCGACCCCACTGCTCAAGAAATTTTGTAAAGCGGACATGTG 960
Qy 961 CCGGTTAGACCATATAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTGAA 1020
Db 961 CCGGTTAGACCATATAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTGAA 1020
Qy 1021 TATACGTACAGTTTACTCCCTTAAACCTGTATGAGATTTTCAAGAGTCTCAAGAT 1080
Db 1021 TATACGTACAGTTTACTCCCTTAAACCTGTATGAGATTTTCAAGAGTCTCAAGAT 1080
Qy 1081 ACTAGCTATTGAAACACTAGTATCGGTGACACCATCATCTCAAGATTTACTAGCT 1140
Db 1081 ACTAGCTATTGAAACACTAGTATCGGTGACACCATCATCTCAAGATTTACTAGCT 1140
Qy 1141 CAAGCAAAAGCATTTTAAACCAAAACCAAGCGGCTATACGATTTTATGAACGTGCTCC 1200
Db 1141 CAAGCAAAAGCATTTTAAACCAAAACCAAGCGGCTATACGATTTTATGAACGTGCTCC 1200

Qy 1201 TCAATCGTCACTCATGACAAATGACATTTTCGTACGATTTTACCAATGATCAAGAGTTT 1260
Db 1201 TCAATCGTCACTCATGACAAATGACATTTTCGTACGATTTTACCAATGATCAAGAGTTT 1260
Qy 1261 ACTTACCGTGTAAAAATCGGGAAACAAAGCTTTATAGGATCAATAAAAAATCTGGTCTGAAT 1320
Db 1261 ACTTACCGTGTAAAAATCGGGAAACAAAGCTTTATAGGATCAATAAAAAATCTGGTCTGAAT 1320
Qy 1321 GAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAGGGGAA 1380
Db 1321 GAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAGGGGAA 1380
Qy 1381 AAGCGTATCATCCCTTTGATCCGAGTCATCTTGAACCTGTTCCACATCAATAGCTTGAAT 1440
Db 1381 AAGCGTATCATCCCTTTGATCCGAGTCATCTTGAACCTGTTCCACATCAATAGCTTGAAT 1440
Qy 1441 GTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAACCTTA 1500
Db 1441 GTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAACCTTA 1500
Qy 1501 GACTTCAGAGATTTATACGATCTCTGATGATAAGGCTAAACTTCTTCAACAATCTCGAT 1560
Db 1501 GACTTCAGAGATTTATACGATCTCTGATGATAAGGCTAAACTTCTTCAACAATCTCGAT 1560
Qy 1561 GCTTTTGGTATTATGGACTATACCTTAACTGGGAAAGTAGAGGATATACGATGACACC 1620
Db 1561 GCTTTTGGTATTATGGACTATACCTTAACTGGGAAAGTAGAGGATATACGATGACACC 1620
Qy 1621 AACCGTATCAATACCGTTTATATGGCAAGCGACCGAAGGAGAGAAATGCTAGCTATCAT 1680
Db 1621 AACCGTATCAATACCGTTTATATGGCAAGCGACCGAAGGAGAGAAATGCTAGCTATCAT 1680
Qy 1681 TTAGCTTATCATAAAGATCGTTATACCGAAGAAAGACGAGAAGTTTACGCTACCTCGGT 1740
Db 1681 TTAGCTTATCATAAAGATCGTTATACCGAAGAAAGACGAGAAGTTTACGCTACCTCGGT 1740
Qy 1741 TATACGGGACACCTATACCTGATACCCCTAAACGACAAATAA 1782
Db 1741 TATACGGGACACCTATACCTGATACCCCTAAACGACAAATAA 1782

RESULT 2

US-10-631-558-12
; Sequence 12, Application US/10631558
; Publication No. US20050260598A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammarra
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/10/631,558
; PRIOR FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US/09/940,235
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2096
; TYPE: DNA
; ORGANISM: Artificial Sequence

```

; FEATURE:
; OTHER INFORMATION: Hybrid cassette
US-10-631-558-12

```

Query Match	94.5%;	Score 1684;	DB 7;	Length 2096;
Best Local Similarity	99.7%;	Pred. No. 0;		
Matches 1687;	Conservative 0;	Mismatches 5;	Indels 0;	Gaps 0;
QY	1	TCGCTTACAGTTCGCTCGCGTATCGGTGATTCATTCCTGCTAAACAGTAAAGCAACCCGCG	60	
DB	51			
QY	61	CAGCTTAGCCGGGTCTCAACGACAGGACGACGATCATGCGCACCCGCTGGCCAGGACCCA	120	
DB	111	CAGCTTAGCCGGGTCTCAACGACAGGACGACGATCATGCGCACCCGCTGGCCAGGACCCA	170	
QY	121	ACGCTGCCGAGATCTCGATCCCGGAAATTAATACGACTCACTATAGGGAGACCAAC	180	
DB	171	ACGCTGCCGAGATCTCGATCCCGGAAATTAATACGACTCACTATAGGGAGACCAAC	230	
QY	181	GGTTTCCTCTAGAAATAATTTTGTTTTAACTTTAAGAAGGAGATATACCATGGTGCAAGC	240	
DB	231	GGTTTCCTCTAGAAATAATTTTGTTTTAACTTTAAGAAGGAGATATACCATGGTGCAAGC	290	
QY	241	ACAAACAGATTGTACCCATAGCTGAGAAAGTGTTTTGATCATGCTGCTGCGACTCTCTATGT	300	
DB	291	ACAAACAGATTGTACCCATAGCTGAGAAAGTGTTTTGATCATGCTGCTGCGACTCTCTATGT	350	
QY	301	GGTGGGAGAAACGTGGGAGAGGACGCGGACGCACTCATTTGCACTTCTAGAAATAGATGC	360	
DB	351	GGTGGGAGAAACGTGGGAGAGGACGCGGACGCACTCATTTGCACTTCTAGAAATAGATGC	410	
QY	361	AACGATCAGGACACACGAGCATCTCTATAGAAATTTGGAGACACCTTGGAGCAAGAAGATAAT	420	
DB	411	AACGATCAGGACACACGAGCATCTCTATAGAAATTTGGAGACACCTTGGAGCAAGAAGATAAT	470	
QY	421	CGAGGAAACCTGCTCCAGTGCATCTGCACAGGCAACGCGGAGAGAGTGGAGTGTGAG	480	
DB	471	CGAGGAAACCTGCTCCAGTGCATCTGCACAGGCAACGCGGAGAGAGTGGAGTGTGAG	530	
QY	481	AGGCACACCTCTGTGCAGACACACATCCAGCGGATCTGGCCCTTTCCCGATGTTTCGTATT	540	
DB	531	AGGCACACCTCTGTGCAGACACACATCCAGCGGATCTGGCCCTTTCCCGATGTTTCGTATT	590	
QY	541	GCTGGACCTGAGTGGCTGTGTAGACCGTCCATCTGTCAACACAGCCAAATTTGTTTATGC	600	
DB	591	GCTGGACCTGAGTGGCTGTGTAGACCGTCCATCTGTCAACACAGCCAAATTTGTTTATGC	650	
QY	601	GTTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTTTGAATCGAT	660	
DB	651	GTTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTTTGAATCGAT	710	
QY	661	CTAAATCATCGACTCTGTATGGAGGAAAGACAGAGCAAGGCTTAAGTCCAAATCAAAA	720	
DB	711	CTAAATCATCGACTCTGTATGGAGGAAAGACAGAGCAAGGCTTAAGTCCAAATCAAAA	770	
QY	721	CCATTTTGCTACTGATAGTGGCGGATGTCAATAAACTTGAGAAAGCTGACTTACTAAAG	780	
DB	771	CCATTTTGCTACTGATAGTGGCGGATGTCAATAAACTTGAGAAAGCTGACTTACTAAAG	830	
QY	781	GCTATTCAAGAACAAATTTGATCGGTAAACGTCACAGTAAACGACGACTACTTTGAGGTCAAT	840	
DB	831	GCTATTCAAGAACAAATTTGATCGGTAAACGTCACAGTAAACGACGACTACTTTGAGGTCAAT	890	
QY	841	GATTTTTCAGAGCGATGCAACCATTTACTGATGCAAAACGGCAAGTCTACTTTGCTGACAAA	900	
DB	891	GATTTTTCAGAGCGATGCAACCATTTACTGATGCAAAACGGCAAGTCTACTTTGCTGACAAA	950	
QY	901	GATGGTTCGGTAAACCTTGGCCGACCAACCTGTCCAGAAATTTTTTGTCTAAGCGACATGTG	960	
DB	951	GATGGTTCGGTAAACCTTGGCCGACCAACCTGTCCAGAAATTTTTTGTCTAAGCGACATGTG	1010	
QY	961	CGCGTTAGACCATATAAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTGAAA	1020	

```

RESULT 3
US-10-631-558-5
Sequence 5, Application US/10631558
Publication No. US20050260598A1
GENERAL INFORMATION:
APPLICANT: Kumar, Rajesh
APPLICANT: Sahni, Rajesh
APPLICANT: Roy, Chaith
APPLICANT: Rajagopal, Kammarra
APPLICANT: Nihalani, Deepak
APPLICANT: Sundaram, Vasudha
APPLICANT: Yadav, Mahavir
TITLE OF INVENTION: NOVEL CLOT-SPEE
TITLE OF INVENTION: PROTEINS POSSI
TITLE OF INVENTION: CHARACTERISTI
TITLE OF INVENTION: PROTEIN
FILE REFERENCE: 07064-009002
CURRENT APPLICATION NUMBER: US/10/
CURRENT FILING DATE: 2003-07-31
PRIOR APPLICATION NUMBER: US/09/94
PRIOR FILING DATE: 2002-04-09

```



```

; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1377
; TYPE: DNA
; ORGANISM: Streptococcus equisimilis
US-10-631-558-5

Query Match          69.9%; Score 1245; DB 7; Length 1377;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 538 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAAATGGTGTGTT 597
DB 133 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAAATGGTGTGTT 192
QY 598 AGCGTTGCTGCTACTCTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAATC 657
DB 193 AGCGTTGCTGCTACTCTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAATC 252
QY 658 GATCTAAACATCAGACCTGCTCATGAGGAGAAACAGACAGCAAGCGCTTAAAGTCCAAATCA 717
DB 253 GATCTAAACATCAGACCTGCTCATGAGGAGAAACAGACAGCAAGCGCTTAAAGTCCAAATCA 312
QY 718 AAACCATTTGCTACTGATAGTGGCGGATGTCAATAAACTTGAGAAAGCTGACTACTA 777
DB 313 AAACCATTTGCTACTGATAGTGGCGGATGTCAATAAACTTGAGAAAGCTGACTACTA 372
QY 778 AAGGCTATTCAAGAAACAATGATCGCTTAAGTCCACAGTCAACAGCACTACTTTGAGTGC 837
DB 373 AAGGCTATTCAAGAAACAATGATCGCTTAAGTCCACAGTCAACAGCACTACTTTGAGTGC 432
QY 838 ATTGATTTTGAAGCGCATGCAACCATTTACTGATCGAAACGGCAAGGTCTACTTTGCTGAC 897
DB 433 ATTGATTTTGAAGCGCATGCAACCATTTACTGATCGAAACGGCAAGGTCTACTTTGCTGAC 492
QY 898 AAGATGTTTGGTAACTTTGCCGACCAACCTGTGCCAAGATTTTGTCTAAGCGGACAT 957
DB 493 AAGATGTTTGGTAACTTTGCCGACCAACCTGTGCCAAGATTTTGTCTAAGCGGACAT 552
QY 958 GTGCGGTTAGACCATATATAAGAAAAACAATACAAAACCAAGCGAAATCTGTGATGTG 1017
DB 553 GTGCGGTTAGACCATATATAAGAAAAACAATACAAAACCAAGCGAAATCTGTGATGTG 612
QY 1018 GAATATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGCCAGGTCTCAA 1077
DB 613 GAATATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGCCAGGTCTCAA 672
QY 1078 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATCTA 1137
DB 673 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATCTA 732
QY 1138 GCTCAAGCAAAAGCAATTTTAAACAAAACCCAGCGGTATACGATTTATGAACGTGAC 1197
DB 733 GCTCAAGCAAAAGCAATTTTAAACAAAACCCAGCGGTATACGATTTATGAACGTGAC 792
QY 1198 TCCTCAATCGTCACTCATGACATGACATTTTCGTTACGATTTTACCAATGGATCAAGAG 1257
DB 793 TCCTCAATCGTCACTCATGACATGACATTTTCGTTACGATTTTACCAATGGATCAAGAG 852
QY 1258 TTACTTACCGTGTAAAAATCGGAAACAGCTTTATAGGATCAATAAAAAATCTGTTCTG 1317
DB 853 TTACTTACCGTGTAAAAATCGGAAACAGCTTTATAGGATCAATAAAAAATCTGTTCTG 912
QY 1318 AATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGCTCTTAAAAAAGGG 1377
DB 913 AATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGCTCTTAAAAAAGGG 972
QY 1378 GAAAAGCCGTATGATCCCTTTTGTATCGAGTCATCTGAACTGTTCACCATCAAAATAGT 1437

```

```

DB 973 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACCTTGAACACTGTTCCACCATCAAAATACGTT 1032
QY 1438 GATGTCGATACCAACGAATTTGCTTAAAGTGAGCAGCTCTTTAAGCTAGCGAAGTAAAC 1497
DB 1033 GATGTCGATACCAACGAATTTGCTTAAAGTGAGCAGCTCTTTAAGCTAGCGAAGTAAAC 1092
QY 1498 TTAGACTTTCAGAGATTTTATACGATCCCTGATAGAGGCTAAACTACTCTTACAAACAATCTC 1557
DB 1093 TTAGACTTTCAGAGATTTTATACGATCCCTGATAGAGGCTAAACTACTCTTACAAACAATCTC 1152
QY 1558 GATGCTTTTGGTATTTATGACTACTATACCTTAACTCGAAAAAGTAGAGGATAATCAGCATGAC 1617
DB 1153 GATGCTTTTGGTATTTATGACTACTATACCTTAACTCGAAAAAGTAGAGGATAATCAGCATGAC 1212
QY 1618 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGGAAGAGAGAATGCTAGCTAT 1677
DB 1213 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGGAAGAGAGAATGCTAGCTAT 1272
QY 1678 CATTTAGCCCTATGATAAGATCGTTATACCGAAGAGAAAGAGAGATTTTACAGCTACTCG 1737
DB 1273 CATTTAGCCCTATGATAAGATCGTTTATACCGAAGAGAAAGAGAGATTTTACAGCTACTCG 1332
QY 1738 CGTTATACAGGACACCTATACCTGTATAACCCCTAACGCAATAA 1782
DB 1333 CGTTATACAGGACACCTATACCTGTATAACCCCTAACGCAATAA 1377

```

```

RESULT 4
US-10-631-558-1
; Sequence 1, Application US/10631558
; Publication No. US20050260598A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammar
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US/10/631,558
; PRIOR FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US/09/940,235
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1245
; TYPE: DNA
; ORGANISM: Streptococcus equisimilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1242)
US-10-631-558-1

```

```

Query Match          69.8%; Score 1243.4; DB 7; Length 1245;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1244; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 538 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAAATGGTGTGTT 597
DB 1 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAAATGGTGTGTT 60
QY 598 AGCGTTGCTGCTACTCTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAATC 657

```

Db 61 AGCGTTGCTGGTACTGTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGGAAATC 120
 Qy 658 GATCTAACATCAGCAGCTGTCTATGAGGAGAAAGACAGAGCAAGGCTTAAAGTCCAAATCA 717
 Db 121 GATCTAACATCAGCAGCTGTCTATGAGGAGAAAGACAGAGCAAGGCTTAAAGTCCAAATCA 180
 Qy 718 AAACCAATTTGCTACTGATAGTGGCGGATGTCTCATAAATTTGAGAAAGCTGACTTACTA 777
 Db 181 AAACCAATTTGCTACTGATAGTGGCGGATGTCTCATAAATTTGAGAAAGCTGACTTACTA 240
 Qy 778 AAGGCTATTCAAGAAATGATGCTTAAGTCCACAGTAACGACGACTACTTTGAGGTC 837
 Db 241 AAGGCTATTCAAGAAATGATGCTTAAGTCCACAGTAACGACGACTACTTTGAGGTC 300
 Qy 838 ATTGATTTTGGCAAGCAGTCAACCATTAATGATGCAAGCGGAGGCTCTACTTTGCTGAC 897
 Db 301 ATTGATTTTGGCAAGCAGTCAACCATTAATGATGCAAGCGGAGGCTCTACTTTGCTGAC 360
 Qy 898 AAGAGTGGTTCGGTAACTTTGCGGACCCAACTGTCCAAGAAATTTTGTAAAGCGGACAT 957
 Db 361 AAGAGTGGTTCGGTAACTTTGCGGACCCAACTGTCCAAGAAATTTTGTAAAGCGGACAT 420
 Qy 958 GTGCGGTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 1017
 Db 421 GTGCGGTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 480
 Qy 1018 GAATATACGTGACGTTTACTCCCTTAAACCCCTGATGACCATTTTACAGCAGGCTCTCAA 1077
 Db 481 GAATATACGTGACGTTTACTCCCTTAAACCCCTGATGACCATTTTACAGCAGGCTCTCAA 540
 Qy 1078 GATCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCAATCTCAAGAAATTA 1137
 Db 541 GATCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCAATCTCAAGAAATTA 600
 Qy 1138 GCTCAAGCAAAAGCAATTTAAACAAACCAAGGCTTATCGATTTATGAAACGTGAC 1197
 Db 601 GCTCAAGCAAAAGCAATTTAAACAAACCAAGGCTTATCGATTTATGAAACGTGAC 660
 Qy 1198 TCCTCAATCGTCACTCATGACATGACATTTTCGTGACATTTTACCAATGGATCAAGAG 1257
 Db 661 TCCTCAATCGTCACTCATGACATGACATTTTCGTGACATTTTACCAATGGATCAAGAG 720
 Qy 1258 TTTTACTACCGTGTAAATAATCGGAAACCAAGCTTATAGGATCAATAAAATCTGGTCTG 1317
 Db 721 TTTTACTACCGTGTAAATAATCGGAAACCAAGCTTATAGGATCAATAAAATCTGGTCTG 780
 Qy 1318 AATGAAGAAATAAACAACTGACCTGATCTCTGAGAAATATTACGCTCTTAAAAAGGG 1377
 Db 781 AATGAAGAAATAAACAACTGACCTGATCTCTGAGAAATATTACGCTCTTAAAAAGGG 840
 Qy 1378 GAAAAGCGGTATGATCCCTTTGATCGAGTCACTTGAATCTTTACCATCAATAGCTT 1437
 Db 841 GAAAAGCGGTATGATCCCTTTGATCGAGTCACTTGAATCTTTACCATCAATAGCTT 900
 Qy 1438 GATGTCGATACCAAGAAATGCTTAAAGTGAGCAGCTTAAACAGCTAGGACGTAAC 1497
 Db 901 GATGTCGATACCAAGAAATGCTTAAAGTGAGCAGCTTAAACAGCTAGGACGTAAC 960
 Qy 1498 TTAGACTTCAGAGATTTATACGATCTCTGTGATGAGGCTAAATCTCTACAAATCTC 1557
 Db 961 TTAGACTTCAGAGATTTATACGATCTCTGTGATGAGGCTAAATCTCTACAAATCTC 1020
 Qy 1558 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAAGTAGAGGATAATCAAGTAC 1617
 Db 1021 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAAGTAGAGGATAATCAAGTAC 1080
 Qy 1618 ACCAACCCTATCATACCGTTTATATGGCAAGCGGACCCGAGAGAGAAATGCTAGCTAT 1677
 Db 1081 ACCAACCCTATCATACCGTTTATATGGCAAGCGGACCCGAGAGAGAAATGCTAGCTAT 1140
 Qy 1678 CATTTAGCCTATGATAAAGATCGTTTATACCGAAGAGAAAGAGGATTTTACAGTACCTG 1737

Db 1141 CATTTAGCTTATGATAAAGATCGTTTATACCGAAGAAACGAGAAAGTTTACAGCTACCTG 1200
 Qy 1738 CGTTATACAGGACACCTTATACCTGATACCTTAACTACGACCAATAA 1782
 Db 1201 CGTTATACAGGACACCTTATACCTGATACCTTAACTACGACCAATAA 1245

RESULT 5
 US-10-631-558-6
 ; Sequence 6, Application US/10631558
 ; Publication No: US20050260598A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kumar, Rajesh
 ; APPLICANT: Sahni, Girish
 ; APPLICANT: Roy, Chait
 ; APPLICANT: Rajagopal, Kammar
 ; APPLICANT: Nihalani, Deepak
 ; APPLICANT: Sundaram, Vasudha
 ; APPLICANT: Yadav, Mahavir
 ; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
 ; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
 ; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
 ; TITLE OF INVENTION: PROTEIN
 ; FILE REFERENCE: 07064-009002
 ; CURRENT APPLICATION NUMBER: US/10/631.558
 ; CURRENT FILING DATE: 2003-07-31
 ; PRIOR APPLICATION NUMBER: US/09/940.235
 ; PRIOR FILING DATE: 2002-04-09
 ; PRIOR APPLICATION NUMBER: 09/471,349
 ; PRIOR FILING DATE: 1999-12-23
 ; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
 ; PRIOR FILING DATE: 1998-12-24
 ; NUMBER OF SEQ ID NOS: 28
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 6
 ; LENGTH: 1327
 ; TYPE: DNA
 ; ORGANISM: Streptococcus equisimilis
 US-10-631-558-6

Query Match 69.1%; Score 1232.2; DB 7; Length 1327;
 Best Local Similarity 99.4%; Pred. No. 0;
 Matches 1237; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 538 ATTCTGCGCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCAATTTGGTGT 597
 Db 83 ATAGCTGGTCTGTAATGGCTACTAGATCGTCTTCTGTAAATAACAGCAATTTGGTGT 142
 Qy 598 AGCGTTGCTGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 657
 Db 143 AGCGTTGCTGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 202
 Qy 658 GATCTAACATCAGCAGCTGTCTGATGAGGAGAAAGACAGAGCAAGGCTTAAAGTCCAAATCA 717
 Db 203 GATCTAACATCAGCAGCTGTCTGATGAGGAGAAAGACAGAGCAAGGCTTAAAGTCCAAATCA 262
 Qy 718 AAACCAATTTGCTACTGATAGTGGCGGATGTCAATAAACTTGAGAAAGCTGACTTACTA 777
 Db 263 AAACCAATTTGCTACTGATAGTGGCGGATGTCAATAAACTTGAGAAAGCTGACTTACTA 322
 Qy 778 AAGGCTATTCAAGAAATTTGATCGCTTAAAGTCCACAGTACGACGACTACTTTGAGGTC 837
 Db 323 AAGGCTATTCAAGAAATTTGATCGCTTAAAGTCCACAGTACGACGACTACTTTGAGGTC 382
 Qy 838 ATTGATTTTGGCAAGCAGTCAACCATTAATGATGCAAGCGGAGGCTCTACTTTGCTGAC 897
 Db 383 ATTGATTTTGGCAAGCAGTCAACCATTAATGATGCAAGCGGAGGCTCTACTTTGCTGAC 442
 Qy 898 AAGAGTGGTTCGGTAACTTTGCGGACCCAACTGTCCAAGAAATTTTGTAAAGCGGACAT 957
 Db 443 AAGAGTGGTTCGGTAACTTTGCGGACCCAACTGTCCAAGAAATTTTGTAAAGCGGACAT 502
 Qy 958 GTGCGGTTAGACCATATAAAGAAACCAATACAAACCAAGCAAGAAATCTGTTGATGTG 1017

Db 503 GTGCGGTTAGACCATATAAGAAAAAACAATACAAAACCAAGCGAATCTGTTGATGTG 562
Qy 1018 GAATATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGTCCTCAA 1077
Db 563 GAATATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGTCCTCAA 622
Qy 1078 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTA 1137
Db 623 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTA 682
Qy 1138 GCTCAAGCAAAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTATGAAACGTGAC 1197
Db 683 GCTCAAGCAAAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTATGAAACGTGAC 742
Qy 1198 TCCTCAATCGTCACTCATGACATGACATTTTCGGTACGATTTTACCAATGGATCAAG 1257
Db 743 TCCTCAATCGTCACTCATGACATGACATTTTCGGTACGATTTTACCAATGGATCAAG 802
Qy 1258 TTTTACTTACCGTGTAAAAATCGGAACAGCTTTATAGGATCAATAAAAAATCTGGTCTG 1317
Db 803 TTTTACTTACCGTGTAAAAATCGGAACAGCTTTATAGGATCAATAAAAAATCTGGTCTG 862
Qy 1318 AATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGG 1377
Db 863 AATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGG 922
Qy 1378 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTACCATCAATACGTT 1437
Db 923 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTACCATCAATACGTT 982
Qy 1438 GATGTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAACAGCTAGCGAACGTAAC 1497
Db 983 GATGTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAACAGCTAGCGAACGTAAC 1042
Qy 1498 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGCTAAACTCTCTACACAAATCTC 1557
Db 1043 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGCTAAACTCTCTACACAAATCTC 1102
Qy 1558 GATGCTTTGGTATTATGACATATACCTTAACTGGAAAAAGTAGAGGATATACGATGAC 1617
Db 1103 GATGCTTTGGTATTATGACATATACCTTAACTGGAAAAAGTAGAGGATATACGATGAC 1162
Qy 1618 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCGAGGAGGATGCTAGCTAT 1677
Db 1163 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCGAGGAGGATGCTAGCTAT 1222
Qy 1678 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAGAAACGAGAAGTTTACAGCTACCTG 1737
Db 1223 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAGAAACGAGAAGTTTACAGCTACCTG 1282
Qy 1738 COTTATACAGGACACTATACCTGATAACCCCTTAACGACAAATAA 1782
Db 1283 CGTTATACAGGACACTATACCTGATAACCCCTTAACGACAAATAA 1327

RESULT 6

US-10-631-558-10
; Sequence 10, Application US/10631558
; Publication No. US20050260598A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammar
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002

; CURRENT APPLICATION NUMBER: US/10/631.558
; CURRENT FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US/09/940.235
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 1661
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hybrid cassette
; US-10-631-558-10

Query Match 64.5%; Score 1150.2; DB 7; Length 1661;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1152, Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 538 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCCAAATGCTGTT 597
Db 183 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCCAAATGCTGTT 242
Qy 598 AGCGTTGCTGTACTGTTGAGGGGACCAATCAAGACATTAAGCTTTAAATTTTTGAAATC 657
Db 243 AGCGTTGCTGTACTGTTGAGGGGACCAATCAAGACATTAAGCTTTAAATTTTTGAAATC 302
Qy 658 GATCTAACATCAACGACCTGCTCATGGAGGAAAGACAGAGCAAGCTTTAAGTCCAAATCA 717
Db 303 GATCTAACATCAACGACCTGCTCATGGAGGAAAGACAGAGCAAGCTTTAAGTCCAAATCA 362
Qy 718 AAACCAATTTGCTACTGATAGTGGCGGATGTACATAAACTTGAGAAAGCTGACTTACTA 777
Db 363 AAACCAATTTGCTACTGATAGTGGCGGATGTACATAAACTTGAGAAAGCTGACTTACTA 422
Qy 778 AAGCTATTCAAGAAACAATTTGATCGCTAAACGTCACAGTAACGACGACTACTTTGAGTGC 837
Db 423 AAGCTATTCAAGAAACAATTTGATCGCTAAACGTCACAGTAACGACGACTACTTTGAGTGC 482
Qy 838 ATTGATTTTCCAGCGATGCAACCATTTACTGATCGAAACCGCAAGGTCCTACTTTGCTGAC 897
Db 483 ATTGATTTTCCAGCGATGCAACCATTTACTGATCGAAACCGCAAGGTCCTACTTTGCTGAC 542
Qy 898 AAAGATGTTTCGGTAACTTTGCGGACCCAACTGTCTCAAGAAATTTTTGCTAAGCGGACAT 957
Db 543 AAAGATGTTTCGGTAACTTTGCGGACCCAACTGTCTCAAGAAATTTTTGCTAAGCGGACAT 602
Qy 958 GTGCGGCTTAGACCATATAAAGAAAAACCAATACAAAAACAGCGAAATCTGTTGATGTG 1017
Db 603 GTGCGGCTTAGACCATATAAAGAAAAACCAATACAAAAACAGCGAAATCTGTTGATGTG 662
Qy 1018 GAATATACCTGACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGTCCTCAA 1077
Db 663 GAATATACCTGACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGTCCTCAA 722
Qy 1078 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTA 1137
Db 723 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTA 782
Qy 1138 GCTCAAGCAAAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTATGAAACGTGAC 1197
Db 783 GCTCAAGCAAAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTATGAAACGTGAC 842
Qy 1198 TCCTCAATCGTCACTCATGACATTAAGCTTTTCGGTACGATTTTACCAATGGATCAAG 1257
Db 843 TCCTCAATCGTCACTCATGACATTAAGCTTTTCGGTACGATTTTACCAATGGATCAAG 902
Qy 1258 TTTTACTTACCGTGTAAAAATCGGGAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 1317
Db 903 TTTTACTTACCGTGTAAAAATCGGGAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 962

```
QY 1318 AATGAAGAAATRAACAACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1377
D 963 AATGAAGAAATRAACAACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1022
QY 1378 GAAAGCCGATGATCCCTTTGATCGCAGTCACTTGAACTGTTTCAACATCAATACGTT 1437
D 1023 GAAAGCCGATGATCCCTTTGATCGCAGTCACTTGAACTGTTTCAACATCAATACGTT 1082
QY 1438 GATGTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 1497
D 1083 GATGTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 1142
QY 1498 TTAGACTTCAGAGATTTATACGATTCCTCGTGATAAGGCTAAACTCTCTACAACAATCTC 1557
D 1143 TTAGACTTCAGAGATTTATACGATTCCTCGTGATAAGGCTAAACTCTCTACAACAATCTC 1202
QY 1558 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAAAAGTAGAGGATAATCAGATGAC 1617
D 1203 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAAAAGTAGAGGATAATCAGATGAC 1262
QY 1618 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAAGGAGAGAAATGCTAGCTAT 1677
D 1263 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAAGGAGAGAAATGCTAGCTAT 1322
QY 1678 CATTTAGCCTATGAT 1692
D 1323 CATTTAGCCGGTGT 1337
```

RESULT 7

```
US-10-631-558-9
; Sequence 9, Application US/10631558
; Publication No. US20050260598A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammarra
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/10/631,558
; CURRENT FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US/09/940,235
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1541
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hybrid cassette
US-10-631-558-9
```

```
Query Match 63.6%; Score 1134.2; DB 7; Length 1541;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1142; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
```

```
QY 538 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCTCAACACGCAATTTGGTTGTT 597
D 33 ATAGCTGGTCCCTGAATGGCTACTAGATCGTCCTTCTGTAAATAACGCCAATTTGGTTGTT 92
```

```
QY 598 AGCGTTGCTGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGTGAATC 657
D 93 AGCGTTGCTGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGTGAATC 152
QY 658 GATCTACATCAGACCTGCTCATGGAGGAAGACAGAGCGGCTTAAGTCCAAATCA 717
D 153 GATCTACATCAGACCTGCTCATGGAGGAAGACAGAGCGGCTTAAGTCCAAATCA 212
QY 718 AAACCATTTGCTACTGATAGTGGGCGGATGTCAATAAACTTGAGAAAGCTGACTTACTA 777
D 213 AAACCATTTGCTACTGATAGTGGGCGGATGTCAATAAACTTGAGAAAGCTGACTTACTA 272
QY 778 AAGGCTATTCAAGAACAAATTGATCGCTAAACGTCACAGTACACGACGACTACTTTGAGGTC 837
D 273 AAGGCTATTCAAGAACAAATTGATCGCTAAACGTCACAGTACACGACGACTACTTTGAGGTC 332
QY 838 ATTGATTTTGCAGCGATGCAACCATTTACTGATCGAAAACGCGAAGGCTACTTTGCTGAC 897
D 333 ATTGATTTTGCAGCGATGCAACCATTTACTGATCGAAAACGCGAAGGCTACTTTGCTGAC 392
QY 898 AAAGATGTTTCGGTAAACCTTGCAGACCCAACTGTCCAAGAAATTTTCTTAAGCGGCAT 957
D 393 AAAGATGTTTCGGTAAACCTTGCAGACCCAACTGTCCAAGAAATTTTCTTAAGCGGCAT 452
QY 958 GTGCGCGTTAGACCATATAAAGAAAAACCAATAACAAAACCAAGCGAAATCTGTTGATGTG 1017
D 453 GTGCGCGTTAGACCATATAAAGAAAAACCAATAACAAAACCAAGCGAAATCTGTTGATGTG 512
QY 1018 GAATATATCTGTACAGTTTACTTCCCTTAAACCCCTGATGACGATTTTCAAGAGGCTCTCAA 1077
D 513 GAATATATCTGTACAGTTTACTTCCCTTAAACCCCTGATGACGATTTTCAAGAGGCTCTCAA 572
QY 1078 GATCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCACTCAAGAAATTTACTA 1137
D 573 GATCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCACTCAAGAAATTTACTA 632
QY 1138 GCTCAAGCAAAAGCATTTTAAACAAAACCAACAGGCTATACGATTTATGAACGTCAC 1197
D 633 GCTCAAGCAAAAGCATTTTAAACAAAACCAACAGGCTATACGATTTATGAACGTCAC 692
QY 1198 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATCGATCAAGAG 1257
D 693 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATCGATCAAGAG 752
QY 1258 TTTACTTACCGTGTAAATAACGGAACAAAGCTTATAGGATCAATAAATACTGGTCTG 1317
D 753 TTTACTTACCGTGTAAATAACGGAACAAAGCTTATAGGATCAATAAATACTGGTCTG 812
QY 1318 AATGAAGAAATRAACAACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1377
D 813 AATGAAGAAATRAACAACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 872
QY 1378 GAAAGCCGATGATCCCTTTGATCGCAGTCACTTGAACTGTTTCAACATCAATACGTT 1437
D 873 GAAAGCCGATGATCCCTTTGATCGCAGTCACTTGAACTGTTTCAACATCAATACGTT 932
QY 1438 GATGTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 1497
D 933 GATGTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 992
QY 1498 TTAGACTTCAGAGATTTATACGATTCCTCGTGATAAGGCTAAACTCTCTACAACAATCTC 1557
D 993 TTAGACTTCAGAGATTTATACGATTCCTCGTGATAAGGCTAAACTCTCTACAACAATCTC 1052
QY 1558 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAAAAGTAGAGGATAATCAGATGAC 1617
D 1053 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAAAAGTAGAGGATAATCAGATGAC 1112
QY 1618 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAAGGAGAGAAATGCTAGCTAT 1677
D 1113 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAAGGAGAGAAATGCTAGCTAT 1172
QY 1678 CATTTAGCCTATGAT 1692
```

```

Db      1173 CATTAGCTGGTGT 1187
|||||
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-105

Query Match      13.1%; Score 232.8; DB 7; Length 2488;
Best Local Similarity 79.8%; Pred. No. 2.3e-59;
Matches 308; Conservative 0; Mismatches 32; Indels 46; Gaps 1;

; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 114
; LENGTH: 2443
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-114

Query Match      13.1%; Score 232.8; DB 7; Length 2443;
Best Local Similarity 79.8%; Pred. No. 2.3e-59;
Matches 308; Conservative 0; Mismatches 32; Indels 46; Gaps 1;

QY      254 CCCATAGCTGAGAAAGTGTGTTGATCATGCTGCTGGGAACTTCTTATGTCGCGAGAAACG 313
          |||||
Db      913 CCCATAGCTGAGAAAGTGTGTTGATCATGCTGCTGGGAACTTCTTATGTCGCGAGAAACG 972
          |||||
QY      314 -----TGGGAGAAAGGCAGC 327
          |||||
Db      973 TGGGAGAAAGCCCTACCAAGCTGGATGATGTTGTTACTTGCCTGGGAGAAAGGCAGC 1032
          |||||
QY      328 GGACGCATCACTTGCACCTTCTAGAAATAGATGCAACGATCAGGACACAAAGGACATCTTAT 387
          |||||
Db      1033 GGACGCATCACTTGCACCTTCTAGAAATAGATGCAACGATCAGGACACAAAGGACATCTTAT 1092
          |||||
QY      388 AGAATTGGAGACACCTGGAGCAAGAGATTAATCGAGAAACCTGCTCCAGTGCATCTGC 447
          |||||
Db      1093 AGAATTGGAGACACCTGGAGCAAGAGATTAATCGAGAAACCTGCTCCAGTGCATCTGC 1152
          |||||
QY      448 ACAGGCAACGCCCGAGGAGAGTGAAGTGTGAGAGGCACACCTCTGTGCAGACCAATCG 507
          |||||
Db      1153 ACAGGCAACGCCCGAGGAGAGTGAAGTGTGAGAGGCACACCTCTGTGCAGACCAATCG 1212
          |||||
QY      508 AGCGGATCTGGCCCTTACCGATGTTGTTGATGCTGAGACCTGATGCTGCTAGACCGT 567
          |||||
Db      1213 AGCGGATCTGGCCCTTACCGATGTTGTTGATGCTGAGACCTGTTTACCAACCGCAGCTCACCCC 1272
          |||||
QY      568 CCATCTGTCAACACAGCCCAATTGGT 593
          |||||
Db      1273 CAGCCTCTCCCTATGCCCACCTGTGT 1298
          |||||

RESULT 10
US-10-995-561-112
; Sequence 112, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 112
; LENGTH: 6510
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-112

Query Match      13.1%; Score 232.8; DB 7; Length 6510;
Best Local Similarity 79.8%; Pred. No. 4e-59;
Matches 308; Conservative 0; Mismatches 32; Indels 46; Gaps 1;

QY      254 CCCATAGCTGAGAAAGTGTGTTGATCATGCTGCTGGGAACTTCTTATGTCGCGAGAAACG 313
          |||||
Db      913 CCCATAGCTGAGAAAGTGTGTTGATCATGCTGCTGGGAACTTCTTATGTCGCGAGAAACG 972
          |||||
QY      314 -----TGGGAGAAAGGCAGC 327
          |||||
Db      973 TGGGAGAAAGCCCTACCAAGCTGGATGATGTTGTTACTTGCCTGGGAGAAAGGCAGC 1032
          |||||
QY      328 GGACGCATCACTTGCACCTTCTAGAAATAGATGCAACGATCAGGACACAAAGGACATCTTAT 387
          |||||
Db      1033 GGACGCATCACTTGCACCTTCTAGAAATAGATGCAACGATCAGGACACAAAGGACATCTTAT 1092
          |||||
QY      388 AGAATTGGAGACACCTGGAGCAAGAGATTAATCGAGAAACCTGCTCCAGTGCATCTGC 447
          |||||
Db      1093 AGAATTGGAGACACCTGGAGCAAGAGATTAATCGAGAAACCTGCTCCAGTGCATCTGC 1152
          |||||
QY      448 ACAGGCAACGCCCGAGGAGAGTGAAGTGTGAGAGGCACACCTCTGTGCAGACCAATCG 507
          |||||
Db      1153 ACAGGCAACGCCCGAGGAGAGTGAAGTGTGAGAGGCACACCTCTGTGCAGACCAATCG 1212
          |||||
QY      508 AGCGGATCTGGCCCTTACCGATGTTGTTGATGCTGAGACCTGATGCTGCTAGACCGT 567
          |||||
Db      1213 AGCGGATCTGGCCCTTACCGATGTTGTTGATGCTGAGACCTGTTTACCAACCGCAGCTCACCCC 1272
          |||||
QY      568 CCATCTGTCAACACAGCCCAATTGGT 593
          |||||
Db      1273 CAGCCTCTCCCTATGCCCACCTGTGT 1298
          |||||

RESULT 9
US-10-995-561-105
; Sequence 105, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 105
; LENGTH: 2488

```

Qy	388	AGAA	TTGGAGACAC	CTGGAGCAAGAGGAT	AATCGAGGAACCTGCTCCAGTGATCTGC	447
Db	1093	AGAA	TTGGAGACAC	CTGGAGCAAGAGGAT	AATCGAGGAACCTGCTCCAGTGATCTGC	1152
Qy	448	ACGG	CAACGGCCGAGAGAGAGT	GGAATGTGAGAGGCACACCTCTGTGCAGACCAACATCG	507	
Db	1153	ACGG	CAACGGCCGAGAGAGAGT	GGAATGTGAGAGGCACACCTCTGTGCAGACCAACATCG	1212	
Qy	508	AGCG	GATCTGGCCCTTTCACCGATGTTTCGTATTCCTGGAACCTGAGTGGCTGTAGACCGT	567		
Db	1213	AGCG	GATCTGGCCCTTTCACCGATGTTTCGTATTCCTGGAACCTGAGTGGCTGTAGACCGT	1272		
Qy	568	CCAT	CTGTCAACAACAGCCCAATTGGT	593		
Db	1273	CAGC	CTCTCCCTATGGCCACTGTGT	1298		

```

RESULT 11
US-10-995-561-117
; Sequence 117, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 117
; LENGTH: 7823
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-117

```

Query Match	13.1%;	Score 232.8;	DB 7;	Length 7823;
Best Local Similarity	79.8%;	Pred. No. 4.4e-59;		
Matches 308;	Conservative 0;	Mismatches 32;	Indels 46;	Gaps 1;
Qy	254	CCCATAGCTGAGAGTGTGTTTGATCATGCTGCTGGGACTTCCTATGTGTGGGAGAAACG	313	
Db	913	CCCATAGCTGAGAAAGTGTGTTTGATCATGCTGCTGGGACTTCCTATGTGTGGGAGAAACG	972	
Qy	314	-----TGGGAGAGGCGCAGC	327	
Db	973	TGGGAGAAAGCCCTACCAAGCGTGCATGATGTTAGTTGATTCCTCGGAGAGGCGCAGC	1032	
Qy	328	GGAGCGATCACTTGCATCTTCAGAAATAGATGCAACGATCAGGACACAGGACATCCTAT	387	
Db	1033	GGAGCGATCACTTGCATCTTCAGAAATAGATGCAACGATCAGGACACAGGACATCCTAT	1092	
Qy	388	AGAATTGGAGACACCTCGAGCAAGAGGATAATCGAGAAACCTGTCTCCAGTGCATCTGC	447	
Db	1093	AGAATTGGAGACACCTGGAGCAAGAGGATAATCGAGAAACCTGTCTCCAGTGCATCTGC	1152	
Qy	448	ACAGGCAACGGCCGAGGAGAGTGGAAGTGTGAGAGGCACACCTCTGTGTGAGACCAATCG	507	
Db	1153	ACAGGCAACGGCCGAGGAGAGTGGAAGTGTGAGAGGCACACCTCTGTGTGAGACCAATCG	1212	
Qy	508	AGCGGATCTGSCCCCTTCCACCGATGTTCTGATTGCTTGGACCTGAGTGGCTGTAGACCGT	567	
Db	1213	AGCGGATCTGSCCCCTTCCACCGATGTTCTGTCAGCTGTTTACCAACCGCAGCCTCACCCC	1272	
Qy	568	CCATCTGTCAACAAACAGCCAAATGGT	593	
Db	1273	CAGCCTCTCCCTATGGCCACTGTGT	1298	

RESULT 12
US-10-995-561-111
; Sequence 111, Application US/10995561

```

; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 111
; LENGTH: 7848
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-111

```

Query Match	13.1%;	Score 232.8;	DB 7;	Length 7848;
Best Local Similarity	79.8%;	Pred. No. 4.4e-59;		
Matches 308;	Conservative 0;	Mismatches 32;	Indels 46;	Gaps 1;
Qy	254	CCCATAGCTGAGAAAGTGTGTTTGGATCATGCTGCGGACATTCCTATGTGTCGAGAAACG	313	
Db	913	CCCATAGCTGAGAAAGTGTGTTTGGATCATGCTGCGGACATTCCTATGTGTCGAGAAACG	972	
Qy	314	-----TGGGAGAAAGCGACG	327	
Db	973	TGGGAGAAAGCCCTACCAAGCTGGATGATGTTGATTGACTTGCCTGGGAGAAAGCGACG	1032	
Qy	328	GGAGCGATCACTTGCATCTTCAGAAATAGATGCAACGATCAGGACACAAAGGACATCCTAT	387	
Db	1033	GGAGCGATCACTTGCATCTTCAGAAATAGATGCAACGATCAGGACACAAAGGACATCCTAT	1092	
Qy	388	AGAAATGGAGACACCTGGAGCAAGAAAGATAATCGAGAAACCTGCTCCAGTGCATCTGC	447	
Db	1093	AGAAATGGAGACACCTGGAGCAAGAAAGATAATCGAGAAACCTGCTCCAGTGCATCTGC	1152	
Qy	448	ACAGGCAACGGCCGAGAGAGTGGAAGTGTGAGAGGCACACCTCTGTGTCAGACCAACATCG	507	
Db	1153	ACAGGCAACGGCCGAGAGAGTGGAAGTGTGAGAGGCACACCTCTGTGTCAGACCAACATCG	1212	
Qy	508	AGCGGATCTGCGCCCTTCACCGATGTTTCGTTATGTTTGGACCTCAGTGGCTGCTAGACCGT	567	
Db	1213	AGCGGATCTGCGCCCTTCACCGATGTTTCGTTATGTTTGGACCTCAGTGGCTGCTAGACCGT	1272	
Qy	568	CCATCTGTCAACACAGCCAAATGGT	593	
Db	1273	CAGCCTCTCCCTATGGCCACTGTGT	1298	

```

RESULT 13
US-10-995-561-113
; Sequence 113, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 113
; LENGTH: 7935
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-113

```

Query Match	13.1%	Score 232.8;	DB 7;	Length 7935;
Best Local Similarity	79.8%;	Pred. No. 4.5e-59;		
Matches 308;	Conservative	0;	Mismatches 32;	Indels 46;
				Gaps 1;

QY 254 CCCATAGCTGAGAAGTGTGTTTGTATGATGCTGCTGGGACTTCCTATGTGGTTCGGAGAAACG 313
Db 913 CCCATAGCTGAGAAGTGTGTTTGTATGATGCTGCTGGGACTTCCTATGTGGTTCGGAGAAACG 972
QY 314 -----TGGGAGAAGGCAGC 327
Db 973 TGGGAGAAGCCCTACCAAGGCTGGATGATGGTAGATTGTACTTGCCTGGGAGAAGGCAGC 1032
QY 328 GGACGCATCATTGCACTTCTAGAAATAGATGCAACGATCAGGACACAAGGACATCCTAT 387
Db 1033 GGACGCATCATTGCACTTCTAGAAATAGATGCAACGATCAGGACACAAGGACATCCTAT 1092
QY 388 AGAATTGGAGACACCTGGAGCAAGAGATAATCGAGAAACCTGCTCCAGTGCATCTGC 447
Db 1093 AGAATTGGAGACACCTGGAGCAAGAGATAATCGAGAAACCTGCTCCAGTGCATCTGC 1152
QY 448 ACAGGCAACGGCGAGAGAGTGGAAAGTGTGAGAGGCACACCTCTGTGCGAGACCACATCG 507
Db 1153 ACAGGCAACGGCGAGAGAGTGGAAAGTGTGAGAGGCACACCTCTGTGCGAGACCACATCG 1212
QY 508 ACCGATCTGGCCCTTACCGATGTTCTGTTTGTGGACCTGAGTGGCTGTAGACCGT 567
Db 1213 ACCGATCTGGCCCTTACCGATGTTCTGTTTGTGGACCTGAGTGGCTGTAGACCGT 1272
QY 568 CCATCTGTCAACAACAGCCAAATTGGT 593
Db 1273 CAGCTCTCTCCCTATGGCCACTGTGT 1298

RESULT 14
US-10-995-561-108
; Sequence 108, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 108
; LENGTH: 7959
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-108

Query Match 13.1%; Score 232.8; DB 7; Length 7959;
Best Local Similarity 79.8%; Pred. No. 4.5e-59;
Matches 308; Conservative 0; Mismatches 32; Indels 46; Gaps 1;
QY 254 CCCATAGCTGAGAAGTGTGTTTGTATGATGCTGCTGGGACTTCCTATGTGGTTCGGAGAAACG 313
Db 913 CCCATAGCTGAGAAGTGTGTTTGTATGATGCTGCTGGGACTTCCTATGTGGTTCGGAGAAACG 972
QY 314 -----TGGGAGAAGGCAGC 327
Db 973 TGGGAGAAGCCCTACCAAGGCTGGATGATGGTAGATTGTACTTGCCTGGGAGAAGGCAGC 1032
QY 328 GGACGCATCATTGCACTTCTAGAAATAGATGCAACGATCAGGACACAAGGACATCCTAT 387
Db 1033 GGACGCATCATTGCACTTCTAGAAATAGATGCAACGATCAGGACACAAGGACATCCTAT 1092
QY 388 AGAATTGGAGACACCTGGAGCAAGAGATAATCGAGAAACCTGCTCCAGTGCATCTGC 447
Db 1093 AGAATTGGAGACACCTGGAGCAAGAGATAATCGAGAAACCTGCTCCAGTGCATCTGC 1152
QY 448 ACAGGCAACGGCGAGAGAGTGGAAAGTGTGAGAGGCACACCTCTGTGCGAGACCACATCG 507
Db 1153 ACAGGCAACGGCGAGAGAGTGGAAAGTGTGAGAGGCACACCTCTGTGCGAGACCACATCG 1212

QY 508 AGCGATCTGGCCCTTACCGATGTTCTGTTTGTGGACCTGAGTGGCTGTAGACCGT 567
Db 1213 AGCGATCTGGCCCTTACCGATGTTCTGTTTGTGGACCTGAGTGGCTGTAGACCGT 1272
QY 568 CCATCTGTCAACAACAGCCAAATTGGT 593
Db 1273 CAGCTCTCTCCCTATGGCCACTGTGT 1298
RESULT 15
US-10-995-561-104
; Sequence 104, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 104
; LENGTH: 8013
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-104

Query Match 13.1%; Score 232.8; DB 7; Length 8013;
Best Local Similarity 79.8%; Pred. No. 4.5e-59;
Matches 308; Conservative 0; Mismatches 32; Indels 46; Gaps 1;
QY 254 CCCATAGCTGAGAAGTGTGTTTGTATGATGCTGCTGGGACTTCCTATGTGGTTCGGAGAAACG 313
Db 913 CCCATAGCTGAGAAGTGTGTTTGTATGATGCTGCTGGGACTTCCTATGTGGTTCGGAGAAACG 972
QY 314 -----TGGGAGAAGGCAGC 327
Db 973 TGGGAGAAGCCCTACCAAGGCTGGATGATGGTAGATTGTACTTGCCTGGGAGAAGGCAGC 1032
QY 328 GGACGCATCATTGCACTTCTAGAAATAGATGCAACGATCAGGACACAAGGACATCCTAT 387
Db 1033 GGACGCATCATTGCACTTCTAGAAATAGATGCAACGATCAGGACACAAGGACATCCTAT 1092
QY 388 AGAATTGGAGACACCTGGAGCAAGAGATAATCGAGAAACCTGCTCCAGTGCATCTGC 447
Db 1093 AGAATTGGAGACACCTGGAGCAAGAGATAATCGAGAAACCTGCTCCAGTGCATCTGC 1152
QY 448 ACAGGCAACGGCGAGAGAGTGGAAAGTGTGAGAGGCACACCTCTGTGCGAGACCACATCG 507
Db 1153 ACAGGCAACGGCGAGAGAGTGGAAAGTGTGAGAGGCACACCTCTGTGCGAGACCACATCG 1212
QY 508 AGCGATCTGGCCCTTACCGATGTTCTGTTTGTGGACCTGAGTGGCTGTAGACCGT 567
Db 1213 AGCGATCTGGCCCTTACCGATGTTCTGTTTGTGGACCTGAGTGGCTGTAGACCGT 1272
QY 568 CCATCTGTCAACAACAGCCAAATTGGT 593
Db 1273 CAGCTCTCTCCCTATGGCCACTGTGT 1298

Search completed: February 1, 2006, 14:43:15
Job time : 331.469 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2006, 16:09:44 ; Search time 7322.56 Seconds
(without alignments)
11385.988 Million cell updates/sec

Title: US-09-940-235-11
Perfect score: 1782
Sequence: 1 tgccttcacgttcgtcgcg.....ataacctaacgacaaataa 1782

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*

- 1: gb_est1.*
- 2: gb_est2.*
- 3: gb_est3.*
- 4: gb_estc.*
- 5: gb_est4.*
- 6: gb_est5.*
- 7: gb_est6.*
- 8: gb_est7.*
- 9: gb_gse1.*
- 10: gb_gse2.*
- 11: gb_gse3.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	233.8	13.1	943	5	BX391752 BX391752
2	232.8	13.1	451	1	AL603368 DKFZp686C
C 3	232.8	13.1	465	1	AI095589 qb24a08.x
4	232.8	13.1	621	7	CN482442 hw20d08.y
5	232.8	13.1	626	7	CN419611 170005313
6	232.8	13.1	765	3	BM715855 UI-E-EJ0-
7	232.8	13.1	7501	4	BC078656 Homo sapi
8	232.8	13.1	7501	4	BC100030 Homo sapi
9	232.8	13.1	7777	4	CR749316 Homo sapi
10	232.8	13.1	7868	4	BM806653
11	232.8	13.1	7885	4	CR749281 Homo sapi
12	232.8	13.1	8411	4	HSMB06992
C 13	231.2	13.0	548	6	CD613789 Homo sapi
14	231.2	13.0	551	6	CD613788 5602208H
15	231.2	13.0	861	1	AU140971 AU140971
16	231.2	13.0	8121	4	CR749317 Homo sapi
17	230.2	12.9	560	7	CN419585 170005318
18	230.2	12.9	737	1	AU140993
19	229.2	12.9	427	7	CN419650
20	229.2	12.9	452	7	CN419649
21	229.2	12.9	547	2	BE009640 PM4-BN017
22	229.2	12.9	580	3	BQ292415 PM0-AN008

23	229.2	12.9	899	5	BX327266
24	227.6	12.8	314	7	CN419631
25	227.6	12.8	496	2	BG945197 PM0-AN008
C 26	227.2	12.7	725	5	BQ574857 UI-H-E21-
27	227	12.7	734	1	AU140802
28	226.6	12.7	517	2	BF081716 PM0-AN008
C 29	225.4	12.6	560	1	AI743013 wg85a12.x
30	223.8	12.6	393	1	AA376374 EST88780
31	219.2	12.3	500	2	BF956982 QV4-NN114
32	217.4	12.2	303	1	AA852090 NHTSCae10
33	216.6	12.2	739	1	AL706215 DKFZp686J
34	215	12.1	705	1	AU140789
35	214.6	12.0	901	5	BX327267
36	214	12.0	598	3	BQ340570 PM0-NN025
37	213.6	12.0	552	2	BG945246 PM0-AN008
C 38	211.6	11.9	564	2	BF229278 PM0-AN008
39	210.8	11.8	603	7	CN419658 170005315
40	210.4	11.8	414	1	AA492032
41	210	11.8	400	3	BQ340575 PM0-NN025
42	209	11.7	577	3	BQ340604 PM0-NN025
C 43	206.8	11.6	367	3	BQ340606 PM0-NN025
44	203.6	11.4	454	8	W46530
C 45	203	11.4	518	1	AI093548 qb08h09.x

ALIGNMENTS

RESULT 1
BX391752/c
LOCUS
DEFINITION BX391752 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA clone CS0D1065Y122 3-PRIME, mRNA sequence.
ACCESSION BX391752 GI:30607701
VERSION BX391752.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 943)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 1413.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna/s=CS08A1025ZG09_CS02375_lkc=1413.r

FEATURES
source Location/Qualifiers
1. .943
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1065Y122"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 13.1%; Score 233.8; DB 5; Length 943;

Best Local Similarity 80.0%; Pred. No. 2.4e-55;
Matches 308; Conservative 0; Mismatches 32; Indels 45; Gaps 1;

Qy 254 CCCATAGCTGAGAAGTGTGTTTGGATCATGCTGCTGGGACTTCCTATGTTGGTCGGAGAAACG 313
Db 778 CCCATAGCTGAGAAGTGTGTTTGGATCATGCTGCTGGGACTTCCTATGTTGGTCGGAGAAACG 719

Qy 314 TG-----GGAGAGGCGAGCG 328
Db 718 TGGGAGAAGCCCTACCAAGCTGGATGGTAGATTGTACTTGGCTGGGAGAAGGCGAGCG 659

Qy 329 GACGCATCATCTTGCACTTCTAGAAATAGATGCAACGATCAGGACACAAGGACATCCTATA 388
Db 658 GACGCATCATCTTGCACTTCTAGAAATAGATGCAACGATCAGGACACAAGGACATCCTATA 599

Qy 389 GAAATTGGAGACACCTGGAGCAGAGGATAATCGAGGAACCTGCTCCAGTCATCTGCA 448
Db 598 GAAATTGGAGACACCTGGAGCAGAGGATAATCGAGGAACCTGCTCCAGTCATCTGCA 539

Qy 449 CAGGCAACCGCCGAGGAGAGTGAAGTGTGAGAGGCACACCTCTGTGCAGACACACATCGA 508
Db 538 CAGGCAACCGCCGAGGAGAGTGAAGTGTGAGAGGCACACCTCTGTGCAGACACACATCGA 479

Qy 509 GCGGATCTGGCCCTTCCACGATGTTCTGATTTGCTGGACCTGAGTGGCTCTAGACCGTC 568
Db 478 GCGGATCTGGCCCTTCCACGATGTTCTGATTTGCTGGACCTGAGTGGCTCTAGACCGTC 419

Qy 569 CATCTGTCTCAACACAGCCCAATTGGT 593
Db 418 AGCCTCTCTCCCTATGCGCCACTGTGT 394

RESULT 2
AL603368
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AL603368
DKFZp686C067.r1.686 (synonym: hlcc3) Homo sapiens cDNA clone
EST.
AL603368.1 GI:15166874
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 451)
This is the 5' sequence of the clone insert
Cloned from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
sequenced by EMBL (European Molecular Biology Laboratories,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No sl sequence available.
This clone (DKFZp686C067) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES
source
1. .451
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686C067"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="686 (synonym: hlcc3)"
/note="Vector: pTriplex2; Site_1: SfiIA; Site_2: SfiIB;
cDNA-collection"

ORIGIN
Query Match 13.1%; Score 232.8; DB 1; Length 451;
Best Local Similarity 79.8%; Pred. No. 3.7e-55;
Matches 308; Conservative 0; Mismatches 32; Indels 46; Gaps 1;

Qy 254 CCCATAGCTGAGAAGTGTGTTTGGATCATGCTGCTGGGACTTCCTATGTTGGTCGGAGAAACG 313
Db 32 CCCATAGCTGAGAAGTGTGTTTGGATCATGCTGCTGGGACTTCCTATGTTGGTCGGAGAAACG 91

Qy 314 -----TGGGAGAAGGCGAGCG 327
Db 92 TGGGAGAAGCCCTACCAAGCTGGATGGTAGATTGTACTTGGCTGGGAGAAGGCGAGCG 151

Qy 328 GGCAGCATCATCTTGCACTTCTAGAAATAGATGCAACGATCAGGACACAAGGACATCCTAT 387
Db 152 GGCAGCATCATCTTGCACTTCTAGAAATAGATGCAACGATCAGGACACAAGGACATCCTAT 211

Qy 388 AGAATTGGAGACACCTGGAGCAAGAGATAATCGAGGAACCTGCTCCAGTCATCTGC 447
Db 212 AGAATTGGAGACACCTGGAGCAAGAGATAATCGAGGAACCTGCTCCAGTCATCTGC 271

Qy 448 ACAGGCAACCGCCGAGGAGAGTGAAGTGTGAGAGGCACACCTCTGTGCAGACACATCG 507
Db 272 ACAGGCAACCGCCGAGGAGAGTGAAGTGTGAGAGGCACACCTCTGTGCAGACACATCG 331

Qy 508 AGCGGATCTGGCCCTTCCACGATGTTCTGATTTGCTGGACCTGAGTGGCTCTAGACCGT 567
Db 332 AGCGGATCTGGCCCTTCCACGATGTTCTGATTTGCTGGACCTGAGTGGCTCTAGACCGT 391

Qy 568 CCATCTGTCAACACAGCCCAATTGGT 593
Db 392 CAGCCTCTCTCCCTATGCGCCACTGTGT 417

RESULT 3
AI095589/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AI095589
qb24a08.x1 Soares_pregnant_uterus_MbHPU Homo sapiens cDNA clone
IMAGE:1697174 3' similar to gb:X02761_cds1_FIBRONECTIN PRECURSOR
(HUMAN); mRNA sequence.
AI095589
AI095589.1 GI:3434565
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 465)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Unpublished (1997)
Tumor Gene Index
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1871 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 446.
Location/Qualifiers
1. .465
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1697174"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Soares_pregnant_uterus_MbHPU"
/note="Organ: uterus; Vector: pT7T3-Pac; Site_1: Not I;
Site_2: Eco RI; 1st strand cDNA was primed with a Not I -

FEATURES
source
1. .465
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1697174"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Soares_pregnant_uterus_MbHPU"
/note="Organ: uterus; Vector: pT7T3-Pac; Site_1: Not I;
Site_2: Eco RI; 1st strand cDNA was primed with a Not I -

Genome Res. 6 (9), 791-806 (1996)

Tissue Procurement: Dr. Gregg Hageman
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).
 Seg primer: M13 Reverse.

```

/clone="UI-E-EJ0-ahj-h-11-0-UI"
/tissue_type="fetal eyes, lens, eye anterior segment,
optic nerve, retina, Retina Foveal and Macular, RPE and
Choroid"
/dev_stage="fetal and adult"
/lab_host="rhl108 (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-EJ0"
/notes="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site 1: Ecor I; Site 2: Not I;
UI-E-EJ0 is a subcloned cDNA library constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1995. First strand cDNA synthesis was primed,
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an Ecor I adaptor, digested
with Not I, and cloned directionally into pT73-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tags for this library are: fetal eyes,
AGATCTCAGA; lens, CGATTACGGA; eye anterior segment,
AATGCCGGAT; optic nerve, CCATTAGTGT; retina, CCGCG; Retina
Foveal and Macular, GTCC; RPE and Choroid, ACCTA. This
library was created for the program, Gene Discovery in the
Visual System, supported by National Eye Institute (NEI)."

```

Query Match 13.1%; Score 232.8; DB 3; Length 765;
Best Local Similarity 79.8%; Pred. No. 4.3e-55;
Matches 308: Conservative 0; Mismatches 32: Indels 46; Gaps 1;

3	QY
2	Db
3	QY
3	Db
3	QY
3	Db
4	QY
4	Db
5	QY

Db 482 AGCGATCTGCGCCCTTACCGATGTTGTCGACGCTGTTTACCAACCGCAGCCTCACCCC 541

Qy 568 CCATCTGTCTCAACACAGCCCAATTGGT 593

Db 542 CAGCCTCTCCCTATGCGCCACTGGT 567

RESULT 7

BC078656

LOCUS

DEFINITION Homo sapiens cdna clone IMAGE:30347017, containing frame-shift errors.

ACCESSION BC078656

VERSION BC078656.1 GI:50925326

KEYWORDS HTC.

SOURCE Homo sapiens

ORGANISM Homo sapiens (human)

REFERENCE

AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Haieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Schetz,T.E., Brownstein,M.J., Usdin,T.B., Tohiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E., Scherch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

CONSRTH

TITLE Mammalian Gene Collection Program Team

human and mouse cdna sequences

Generation and initial analysis of more than 15,000 full-length human and mouse cdna sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 7501)

NIH MGC Project

Direct Submission

Submitted (02-AUG-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgabs-x@mail.nih.gov

Tissue Procurement: Dr. Stefan Hansson

cdna Library Preparation: Michael Brownstein / Ted Usdin

Laboratory

cdna Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www-shgc.stanford.edu>

Contact: (Dickson, Mark) mcd@paxil.stanford.edu

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>

Series: IRAC Plate: 168 Row: 0 Column: 20

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 47132556

This clone has the following problem: frame shifted.

Location/Qualifiers

1. .7501

/organism="Homo sapiens"

FEATURES

source

Query Match 13.1%; Score 232.8; DB 4; Length 7501;

Best Local Similarity 79.8%; Pred. No. 8.6e-55;

Matches 308; Conservative 0; Mismatches 32; Indels 46; Gaps 1;

Qy 254 CCCATAGCTGAGAAAGTGTTCATCTGCTGGGACTTCTCTATGTGGTGGGAGAAACG 313

Db 806 CCCATAGCTGAGAAAGTGTTCATCTGCTGGGACTTCTCTATGTGGTGGGAGAAACG 865

Qy 314 -----TGGGAGAAAGGCAGC 327

Db TGGGAGAAAGCCCTACCAAGGCTGGATGATGTAGATTGTACTTGGCTGGGAGAGGCAGC 925

Qy 328 GGAGCGCATCTTGCACCTTCTAGAAATAGATGCAACGATCAGGACACAGGACATCCTAT 387

Db 926 GGAGCGCATCTTGCACCTTCTAGAAATAGATGCAACGATCAGGACACAGGACATCCTAT 985

Qy 388 AGAATTGGAGACACCTGGAGCAAGAGGATTAATCGAGGAAACCTGCTCCAGTGCATCTGC 447

Db 986 AGAATTGGAGACACCTGGAGCAAGAGGATTAATCGAGGAAACCTGCTCCAGTGCATCTGC 1045

Qy 448 ACAGGCAACGGCCGAGGAGAGTGGAAAGTGTGAGAGGCACACCTCTGTGACAGCACATCG 507

Db 1046 ACAGGCAACGGCCGAGGAGAGTGGAAAGTGTGAGAGGCACACCTCTGTGACAGCACATCG 1105

Qy 508 AGCGGATCTGCGCCCTTCCACCGATGTTGTTATGCTGGACCTGAGTGGCTGCTGACAGCT 567

Db 1106 AGCGGATCTGCGCCCTTCCACCGATGTTGTTATGCTGGACCTGTTTACCAACCGCAGCCTCACCCC 1165

Qy 568 CCATCTGTCAACAACAGCAGCAATTGGT 593

Db 1166 CAGCCTCTCTCCCTATGGCCACTGTGT 1191

RESULT 8

BC100030

LOCUS

DEFINITION Homo sapiens cdna clone IMAGE:30343682, containing frame-shift errors.

ACCESSION BC100030

VERSION BC100030.1 GI:71679847

KEYWORDS HTC.

SOURCE Homo sapiens

ORGANISM Homo sapiens (human)

REFERENCE

AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Haieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Schetz,T.E., Brownstein,M.J., Usdin,T.B., Tohiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:30347017"

/tissue_type="Placenta, normal"

/clone_lib="NIH MGC_147"

/lab_host="DH10B"

/note="vector: pBluescriptR"

Schnarch, A., Schein, J.E., Jones, S.J. and Warra, M.A.
Mammalian Gene Collection Program Team
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 7501)
NIH MGC Project
Direct Submission
Submitted (01-AUG-2005) National Institutes of Health, Mammalian
Gene Collection (MGC), Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-x@mail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
cDNA Library Preparation: Michael Brownstein / Ted Usdin
Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 198 Row: j Column: 21
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 47132556
This clone has the following problem: frame shifted.
FEATURES
source
1..7501
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30343682"
/tissue_type="Placenta, normal"
/clone_lib="NIH_MGC_147"
/lab_host="DH10B"
/note="vector: pbluescriptR"
ORIGIN
Query Match 13.1%; Score 232.8; DB 4; Length 7501;
Best Local Similarity 79.8%; Pred. No. 8.6e-55;
Matches 308; Conservative 0; Mismatches 32; Indels 46; Gaps 1;
QY 254 CCCATAGCTGAGAAGTGTCTTTTGATCATGCTGCTGGGACTTCCTATGTGGTCGGAGAAACG 313
Db 806 CCCATAGCTGAGAAGTGTCTTTTGATCATGCTGCTGGGACTTCCTATGTGGTCGGAGAAACG 865
QY 314 -----TGGAGAAAGGAGC 327
Db 866 TGGAGAAAGCCCTACCAAGGCTGGATGATGGTAGATTGTACTTTCCTCGGAGAAAGGAGC 925
QY 328 GGAAGCATCACTTGCATCTCTAGAAATAGATGCAACGATCAGGACACAAGACATCCTAT 387
Db 926 GGAAGCATCACTTGCATCTCTAGAAATAGATGCAACGATCAGGACACAAGACATCCTAT 985
QY 388 AGAATTGGAGACACCTGGAGCAAGAGGATAATCGAGGAAACCTGTCTCCAGTGCATCTGC 447
Db 986 AGAATTGGAGACACCTGGAGCAAGAGGATAATCGAGGAAACCTGTCTCCAGTGCATCTGC 1045
QY 448 ACAGGCACCGCCGAGAGAGTGGAAGTGTGAGAGGCACACCTCTGTGCGAGACCAATCG 507
Db 1046 ACAGGCACCGCCGAGAGAGTGGAAGTGTGAGAGGCACACCTCTGTGCGAGACCAATCG 1105
QY 508 AGCGGATCTGC3CCCTTTCACCGATGTTTCGATTGCTCGACCTGAGTGGCTGTAGACCGT 567
Db 1106 AGCGGATCTGC3CCCTTTCACCGATGTTTCGATTGCTCGACCTGTTTACCAACCGCAGCCTCACCCC 1165
QY 568 CCATCTGTCAACACACGACCAATTGGT 593

Db 1166 CAGCCTCTCCCTATGCGCACTGCT 1191

RESULT 9
 CR749316
 LOCUS
 DEFINITION Homo sapiens mRNA; cdna DKFZp686K08164 (from clone DKFZp686K08164).
 ACCESSION CR749316
 VERSION CR749316.1 GI:51476361
 KEYWORDS HTC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 7777)
 Koehrer,K., Beyer,A., Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and Wiemann,S.
 The German CDNA Consortium
 Direct Submission
 Submitted (17-AUG-2004) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by BMFZ (Biomedical Research Center at the Heinrich-Heine-University, Duesseldorf/Germany) within the CDNA sequencing consortium of the German Genome Project. This clone (DKFZp686K08164) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering:
 http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp686K08164
 Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.
 Location/Qualifiers
 1..7777
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="DKFZp686K08164"
 /tissue_type="uterus endothel, primary cell culture"
 /clone_lib="686 (synonym: hlcc3). Vector pSport1_sfi; host DH10B; sites SfiIa + SfiIB"
 /dev_stage="adult"
 /note="fibronectin 1, differentially spliced"
 1..7777
 /gene="DKFZp686K08164"
 266..7069
 /gene="DKFZp686K08164"
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="CAH18171.1"
 /db_xref="GI:51476362"
 /translation="MLRGPQGLLLLAVLCLGTAVPSTGASKRRQAQVMQVQSPVA
 VNSQKFGYDNGHYQINQWERTYILGNALVCTCYGSGRGNCEKPEASETCFDKTY
 GNTYRGVDTYRHPKSWINDCTCIGAGRGISICTIANRCHGGGYSYKIGDTWRPRPST
 GGYMECVCLGNCKGKSWTCXPIAEKCFDHAAGTSYVVGTEWKPYPQCGMMVADCTCLGE
 GSRICTTSNRNCNDQTRTSYRIAGDTWSKKDNRGNLLQICITGNGSGEKKCRHSTVS
 QITSSGSCPTDRAAIVYQPPHPPPYGHCVCTDSCGVYSVGMQWLKTKGNQKMLCT
 CLNGVSCQETAVTQYTGSGNSGAPCVLPFTYNGNRTFYSCTTEGRQDHLWCSTTSNY
 EDOQTSFCETHTVLTQYTGSGNSGALCHFPFYLRNNHYTDCSTSEGRDNMKRGWETC
 NYADQKFGFPGMAAHEICVTINNEGVMYRIQDQWKDQHDGMHMRCTCGVNGGEWTC
 IAYSQRLDQCIDVDITVNNVDTHKREHGGMLNCTCFGQGRGWKCDPVDQDQDSST
 GTFYQIGDSWEKYVHGVQCYCYGRCIGEHWCQPLQTYPSSSGPVEFVETEPQEN
 SHPIQWNAQPSHSIKYLKWRPNKSVGRKWEATIPGLNLSYTIKGLKPGVWVEGQLI
 SIQQYGHQEVREDFITPTTSNTVNTGTTPSPVATSESVTEITASSFVSQW
 VASLSTTSOTPAPEYELSEEGDEPQYLDLPSTATSVNIPDLLPKRKYIVNVYISDEGE
 QSDITSTFADPDPDPTVDQDDTSTVVRWSRQPAITGYRIVYSPSVSGSSTEL
 NLPETANSVTLSDLPQGVQNTITYAEEQESTPEVVIQETTTGTPSTDPVSPRDLQ
 PFVTDVKYTIMTPPESATGYRVDVIVNLPGEHQRLPISRNTAEVTVGLSPGVY
 YTPKFAVSHGSKSLQAQITKLDAPTNLQFVNEDTSIVLRWTPPRAQITGYRIL
 VSGTRQPRQYNGVSVSKYPLRNLPQASEYTVSLVAIKNGQESPRATGVTTLQPG
 SSIPTNYNTEVTITVITWTPAPRIKFLGRVPSQGEAREVTSDSGSIWVSGLTGLT

VEVVTIOVLDRQERDAPVKNVVTPLSPPTNLHLEANPDVTGLTVSWERSTTPDIT
GVRITITPQNGNSLEEVHADQSSCTFDNLSPLGLEYNVSVTVVTKDKESFVDSI
IPEVPLDLSFVDSISGLRWTPLNSSTIIGRITVVAAGEGPIPEDFVDSV
GYTVTGLBPGIDISVITPLNGESAPTLTQQTAVPPTDLFTNIGDNRVVTW
APPSIDLFLVYSPVKNEDVAELISPSDNAVLTLNLPGLTEVVSWSVVEOH
ESTPLGRKOTGLDSTGIDFSDITANSFVHIAIPRATITGVRIHHPHFSGRPRE
DRVPHRSNITLNLPTGYEYVSVIALNGRRESPLLIQQOSTVSDVPRDLEVVAATP
TSLIISWAPDVTIRITGETGNSPVQEFVPGSKSTATISGLKPGVDYITVY
AVNGEDSPASSKPIINVTREIDKPSQMQVTDVQNSISVKWLPSSPVTYRVITP
KPGPGPTKTKTAGPQTEMTIECLOPTEYVYVVAQNPSSGSOPLVQVATVTPAP
TDLTKVQPTLSAOWTPNVOLTVRVVTPKEKTPMKENLAPDSSSVVGSLM
VATKVEVVALKDTITSRPAQGVTTLENVSPRARVTDATETITISWRKTEFI
TGFQDAPVANGCPTQRTIKPVRYSYITGLQDYKILYTLNDNARSPPVIDA
STAIAPSNLEFIATVYVNSVQPPRATITGYIKYKPGSPPRVPRPGRVTE
ATITGLEPTEYVIALKNKQSEPLIGRKTQGEALQSTTISWAPFQDTEYIIS
CHPVGDPTSTVSTVATLGLTRGATNIIIEALKDQORHKVREEVTVGNS
VNEGLNPTDSDCFDPTVSHYAVDEWERMSSGPKLLCQCLGRSGHRCDSRWC
HONGVNYKIGEKWDROGQMSCTCLNGKGEPCDPRHATCYDDKTYHYGEQWQ
KEYLGAICSCCTFGQGRWRCNRRPGEPTTGGTGGQINQTSQKHQRTNIVNC
PIECFMPLDVQADRESRE"

ORIGIN

Query Match	13.1%	Score	232.8	DB	4	Length	7777
Best Local Similarity	79.8%	Pred.	No. 8.7e-55				
Matches	308	Conservative	0	Mismatches	32	Indels	46
Gaps	1						
QY	254	CCCATAGCTGAGAAGTGTGTTGATCATCTGCTGGGACTTCTCTATGTTGGTCGGAAGAACG	313				
Db	806	CCCATAGCTGAGAAGTGTGTTGATCATCTGCTGGGACTTCTCTATGTTGGTCGGAAGAACG	865				
QY	314	-----TGGGAGAGGCGAGC	327				
Db	866	TGGGAGAGGCGCTACCAAGCTGGATGATGTTAGATTGTTACTTGCCTGGGAGGAGCGAGC	925				
QY	328	GAAGCATCACTGCACTTCTAGAAATAGATGCAACGATCAGGACACAGCATCTCTAT	387				
Db	926	GAAGCATCACTTGGCACTTCTAGAAATAGATGCAACGATCAGGACACAGCATCTCTAT	985				
QY	388	AGAAATTGGAGACACCTGGAGCAAGAGGATAATCGAGGAACCTGCTCCAGTGATCTGC	447				
Db	986	AGAAATTGGAGACACCTGGAGCAAGAGGATAATCGAGGAACCTGCTCCAGTGATCTGC	1045				
QY	448	ACAGGCAACGGCGGAGGAGAGTGAAGTGTGAGAGGCACACCTCTGTGCAGACCAACATCG	507				
Db	1046	ACAGGCAACGGCGGAGGAGTGAAGTGTGAGAGGCACACCTCTGTGCAGACCAACATCG	1105				
QY	508	ACGGATCTGGCCCTTACCGATGTTGTTATGTTGAGACCTGAGTGCGTGTAGACCGT	567				
Db	1106	ACGGATCTGGCCCTTACCGATGTTGTTATGTTGAGACCTGAGTGCGTGTAGACCGT	1165				
QY	568	CCATCTGTCAACACAGCCAATTGGT	593				
Db	1166	GAGCTCTCTCCCTATGGCCACTGTGT	1191				

RESULT 10	HSM806653	7868 bp	mRNA	linear	HTC 20-JAN-2005
LOCUS	Homo sapiens mRNA; cDNA DKF2p686M04163 (from clone DKF2p686M04163).				
DEFINITION	BX640608				
ACCESSION	BX640608.1	GI:34364616			
VERSION	HTC.				
KEYWORDS	Homo sapiens (human)				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 7868)				
AUTHORS	Bloeker, H., Boecher, M., Brandt, P., Mewes, H.W., Weil, B., Amid, C., Osanger, A., Fobo, G., Han, M. and Wiemann, S.				
CONSRSTM	The German cDNA Consortium				
TITLE	Direct Submission				
JOURNAL	Submitted (20-JAN-2005) MIPS, Ingolstaedter Landstr.1, D-85764				

COMMENT

Neuherberg, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@kfz-heidelberg.de;
sequenced by GBF (National Research Centre for Biotechnology Ltd., Braunschweig/Germany) within the cDNA sequencing consortium of the German Genome Project.
This clone (DKF2p686M04163) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKF2p686M04163
Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.

FEATURES

location/Qualifiers
1..7868
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RZPD:DKF2p686M04163Q"
/db_xref="taxon:9606"
/clone="DKF2p686M04163"
/tissue_type="uterus endothel, primary cell culture"
/clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host DH10B; sites SfiI + SfiIb"
/dev_stages="adult"
/note="fibronectin precursor"
1..7868
/gene="DKF2p686M04163"
267..7157
/gene="DKF2p686M04163"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAE45714.1"
/db_xref="GI:34364617"
/db_xref="GOA:Q6N0A6"
/db_xref="InterPro:IPR000083"
/db_xref="InterPro:IPR000562"
/db_xref="InterPro:IPR002086"
/db_xref="InterPro:IPR003961"
/db_xref="InterPro:IPR006209"
/db_xref="UniProt/T-EMBL:Q6N0A6"

/translation="MLRGPGLLLALCLGTAVPSTGASKSRQKQKQVQPSVPA
GSKPGCYDNGHYQINQWERTYLGALVYCTCYGRGFCNCKEPEAEETCFDKYT
VNTYRGDTYERKDSIMIDCTCIGAGRGVICTIANRCHGGQSYKIGDTWRPHET
VMSLEVCCLNGKGEWTKPIAEKCFDHAAGTSYVVGTEWKEPYQGMVVDCTCLG
GSGRICTSRNRCNDQRTSYRIIGDTWSKDNGLLQICITCNGHGEKCEHRTSV
QTSSGSGPTDRAAYIQPHPPPPYCHVCTDVGTVSGVVGWMLKTKQNKMLCT
ELQNGVSCQETAVTQITGGNSGEPCLVLPFTYNGRTFYSCITTEGRQGLHWCSTSY
CQQRYSFCSTHTVLVQTRGNSGALCHFPFLYNNHYDCTSEGRDNMKWCTTQ
IAYSQLDQCIQVDDITVYVNDTFHKEEGHMLNCTCGGGRGWKCDPVDQCDSET
GTFVQIGDSMEKYVHYVQCYCGRGIGWHCQPLQTYPSSSGPVFVFTETSPQN
SHIQNAPQPSHISKILRWPRKNSVGRKEATIPGLNSYITIGLKPVGTVYEGQLI
SIQYGHQEVTRFDFTTSTPTVNTVTGTTTTPPLVATSESTVEITASSFVSM
VSASDTVGFVREYELSEEGEPQYLDLPSTATSVNIPDLIPGRKYTVNYQISEDE
QSLISTQTTAPADPDVDVDTSIIVVRSRQAPITGVRIYVSPSVGSGSTEL
NLPTANSVTLSDLRQGVQVNIITIVAEVNOESTPVVIQOETGTGTPSDTVPSPRDQ
FVEVDVKTIMMTTPESAATGYRVDVIVNLPGEHQRLPISENTEAETVGLSPGT
IYFKFAVSHGRSKPLTAQTTKLDAPTLQFVNEDTSLVWMTPTPRAQITGYRLT
VGLTRRGQRYNVGSPVSKYPLRNLQPSAEYTVSLVIAIKNQESKATGTVLTPQ
SSIPYNTTEVTTITVTTPAPRIGKLVGSPSQGEAPREVTSDSGTVVSLTPG
VEVYITQVLRDQERDAPIVNKVVTPLSPNLHLEANPDVTGLTVSWERSTTPDIT
GVRITITPQNGNSLEEVHADQSSCTFDNLSPLGLEYNVSVTVVTKDKESFVDSI
IIPAVPPDRLFTNIGDPTMRTVMAFPSSIDLTLNPLVRSYVKNEDVAELISPSD
NAVLTNLLPGTEYVSVSVSQEHESTPLRGKQTLGSDTDFDFTNLSFTVHM
IAPRATITGYRIHHPHFSGRPREDRPHRSNITLNLPTGYEYVSVIALNGRE
SPVLIQQOSTVSDVPRDLEVVAATPSTSLIISWAPVTVVYRIYTGTCGNSQVQF
TVPGSKSTATISGLKPGVDYITVYVTVGSDSPASSKPIINVTREIDKPSQMVDT
VQNSISVKWLPSSPVTYRVITPKNPGPPTKTAGPDQTEMTIEGLQPTVEYV
VTAQNPSSGSOPLVQVATVITIPAPDLKFTQVTTLSLAQMTTPPNVQLTGYRVTP
KEKTPMKENLAPDSSSVVGSLMVAATKVEVVALKDTITSRPAQGVTTLENVSP
PRARVTDATETITISWRKTEFI
PGTDYKLYTLNDNARSPPVIDA
TDAIDAPSNLFLATTNLSLIDKVDVSTYITGLQ
YIYKPKGSPREVPRPGRVTEATITGLEPCTEYTIYVIALKNKQSEPLIGRKT
TDLPLQLVTLPHNLHGPILDVPSTVQKTPFTVTHPGDTGNGIQLPGTSQQQPSVQ

Db	966	GGAGCGATCACTTGGCACTTCTAGAAATAGATGCAACGATCAGGACACAAAGGACATCTTAT	1025
Qy	388	AGAATTGGAGACACCTGGAGCAAGAGGATAATCGAGAAACCTGCTCCAGTGCATCGC	447
Db	1026	AGAATTGGAGACACCTGGAGCAAGAGGATAATCGAGAAACCTGCTCCAGTGCATCGC	1085
Qy	448	ACAGGCAACGGCCGAGAGAGTGAAGTGTGAGAGGACACCTCTGTGACAGACCATCG	507
Db	1086	ACAGGCAACGGCCGAGAGAGTGAAGTGTGAGAGGACACCTCTGTGACAGACCATCG	1145
Qy	508	ACGGGATCTGGCCCTTCCAGATGTTCTGATTTGCTGCACTGAGTGGCTGTAGACGT	567
Db	1146	ACGGGATCTGGCCCTTCCAGATGTTCTGATTTGCTGCACTGAGTGGCTGTAGACGT	1205
Qy	568	CCATCTGTCAACACAGCCAAATGGT	593
Db	1206	CAGCTCTCTCCCTATGGCCACTGTGT	1231

RESULT 12
HSM806992
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
HSM806992 8411 bp mRNA linear HTC 23-FEB-2005
Homo sapiens mRNA; cDNA DKFZp686O1166 (from clone DKFZp686O1166).
BX640875
BX640875.1 GI:34365169
HTC.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1 (bases 1 to 8411)
Bloeker, H., Boecher, M., Brandt, P., Mewes, H.W., Weil, B., Amid, C.,
Osanger, A., Fobo, G., Han, M. and Wiemann, S.
The German cDNA Consortium
Direct Submission
Submitted (22-FEB-2005) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by GBF (National Research Centre for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.
This clone (DKFZp686O1166) is available at the RZPD Deutsches
Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp686O1166
Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/.

FEATURES
source
Location/Qualifiers
1..8411
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RZPD:DKFZp686O1166Q"
/db_xref="taxon:9606"
/clone="DKFZp686O1166"
/tissue_type="endometrium carcinoma cell line"
/clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host
DH10B; sites SfiIA + SfiIB"
/dev_stage="adult"
/note="fibronectin 1"
1..8411
/gene="DKFZp686O1166"
268..7701
/gene="DKFZp686O1166"
/codon_start=1
/product="hypothetical protein"
/protein_id="CA545932.1"
/db_xref="GI:34365170"
/db_xref="GOA:Q6MZU5"
/db_xref="InterPro:IPR000083"
/db_xref="InterPro:IPR000562"

gene
CDS

/db_xref="InterPro:IPR002086"
/db_xref="InterPro:IPR003961"
/db_xref="InterPro:IPR006209"
/db_xref="UniProt/TREMBL:Q6MZU5"
/translation="MLRPGPGLLLLLAVLCIGTAVPSTGASKSKRQAQOMVQPSFVA
VSKPGCTDNGKHQIINQWERTYLLNALVCTCYGSGRGNFCEKSPABETCFDKYT
GNTVRGDTYERPKDSMTDCTCIGAGRISCTIANRCHGGQSYKIGDWRPHET
GMYLCEVCLNGKEWTKCPAIEKCFDHAAGTSYVVGSTWKEPTQGMWMDCTCLGE
GSRITCTSRNRKNDQDRTSVYIGDTSKKDNRLNLCQICTNRGRGECERTSV
QTSSTSGSPFTDRAAVYQPPHPPPPYCHVTDSGVVYVGMQMLCTQGNQMLCT
CLNGVSCQETAVTQYCGNSNGEPCVLPFTYNGRTFYSCTTEGRQDGLHWCSTTSY
EQDKISFCTDHTLVLRGNSNGALCHFPFLYNNHNTDCTSGRDNMKWCCGTQ
NYDADQKQFCHMAAHEICTTNGMYRIGDQMDQMDHMMRCTCVGNGSGWTC
IAYSLQRDQCIDVDTYNNDTFKRHEEGLNCTCFQGRGRKCDPVDQDSPT
GTFYIGDSWEKYVHGVYQVCYGRGIGEWHCPLQTPSSSGPVEVPIETPSQPN
SHPLQWNAPOPSHLSKYLRLWRPKNSVGRKEATIPGLNSVTIKGLKPGVVEQLI
SIQYGHQEVTRFPTTSTPSTNSVTGETTFFPSLVAISEVSTIASSFVSW
VSADTVSGFRVEYELSEBDEPQYLDLPSTATSVNIPLLLPGRKYIVNVQISDGE
QSLILSTSQTAPADPPTVDVDDTSIVVMSRPOAPITGTYRIYVPSVSGSTEL
NLPTANSVTLSDLPQGVQYNIIVAVEENOESTPVIIQOETGTPRSDTVPSRDLQ
FVEYTDVKVTIMWTPPEASATGYRVDVLPVLPGEHGORLPISRTFAEVTGLSPGT
YFKFVAFVSHGRESKPLTAQOTTKLDAPTNLQFVNETSVLVRWTPPRAQITGLRLT
VGLTRGQPRQNVGPSVSKYLRNLQPASETYVSLVAIKGNQSPKATGPTTLPQG
SSIPPYNTEVTETTVITPAPRIGFLGVRPSQSGAPREVTSDSGSVVSGLTGPG
VEYVYITQVLRDQGERDAPIVKVVPLSPPTNLHLNAPDGTGLVTSWERSITPDIT
GYRITPTPTNGQOQNSLEEVQADQSSCTFDNLSPGLEVNVVTVKDDKESVPLSDT
IIEVPOLTDLSFVDIITDSSIGLWTPPLNSSTIIGYRITVVAAGEGIPIDFDFVDSV
GYTVTGLEPGIDVIDSVITLINGSAPTTTQQTAPPTPPTDLPRTNIGDPTMKVTW
PANGFTKTGAPDQETEMTIEGLQFVEYVSVIAQNSPESQPLVQTAVTNDLRP
KGLATDVDDSIKIAMESQGVSRVRYTVSSPEDGHELFAPDGEDFAELQGLR
PGEYTVSVVALHDMESQPLIGTQSTAPAPTDLKTQVTVTSLSAQWTPNVQLGT
YRVVTPKEKTGPMKEINLAPDSSVSVGLTVATKYEVSVALKDTSRPAQCVLT
TLENVSPRRARVTDATETTTISWRTTETITGQVDAPANGOTPIORTIKPDVRS
TYLNGLPQTDYKIYLTINDNARSPPVIDASTALDAPSLRFLATPNSLLVSWQP
PRATIGYIIKIKRPGSPREVPRPRPTEATITGLEPGETYIYVIAUKNNQKSE
PLIKRKTDELQPLTLPHPNLHGEILDVSTVQKTFVTPGDTGNGTQLRGTSG
QPSVGQGMPTPEEHGFRRTPTPTPIRHRPRYPNPVGEIIOIHPREDVDVHLY
PHGPLNASTGOEALSOITISWAPFOTSEYIISCHPVGTDEEPLOFRVPGTSTA
TLGLTRGATYNIIVEALKOORHKVREEVTVGNSVNEGLNQITDDSCFPDPTVSHY
AVGEWERSSESGFKLJCCLGSGHFRCDSSRWCHDNGVNIKGEKWDGNGQM
MSTCLGNGKEGFECDPHEATCYDDGTYHVGEQWQKYLGAICSTCFGQGRWCRD
NCRPPGGEPSPEGTTGQSYNQYSQRYHQRTNTVNYPIECFMPDLVQADREDSRE"

ORIGIN
Query Match 13.1%; Score 232.8; DB 4; Length 8411;
Best Local Similarity 79.8%; Pred. No. 8.9e-55;
Matches 308; Conservative 0; Mismatches 32; Indels 46; Gaps 1;
Qy 254 CCCATAGCTGAGAAAGTGTGTTTGTATCATGCTGCTGGAGCTTCTCTATGTGTCGCGAGAAACG 313
Db 808 CCCATAGCTGAGAAAGTGTGTTTGTATCATGCTGCTGGAGCTTCTCTATGTGTCGCGAGAAACG 867
Qy 314 -----TGGGAGAAAGCGCAGC 327
Db 868 TGGGAGAAAGCCCTACCAAGCTGGATGATGCTAGATTGTACTTGGCTGGGAGAGGCGAGC 927
Qy 328 GGACGCACTCACTTGCACCTTCTAGAAATAGATGCAACGATCAGGACACAGGACATCTTAT 387
Db 928 GGACGCACTCACTTGCACCTTCTAGAAATAGATGCAACGATCAGGACACAGGACATCTTAT 987
Qy 388 AGAATTGGAGACACCTGGAGCAAGAGATATCGAGGAAACCTGCTCCAGTGCATCTGC 447
Db 988 AGAATTGGAGACACCTGGAGCAAGAGATATCGAGGAAACCTGCTCCAGTGCATCTGC 1047
Qy 448 ACAGGCAACGGCCGAGAGAGTGAAGTGTGAGAGGACACACTCTGTGTCAGACACCATCG 507
Db 1048 ACAGGCAACGGCCGAGAGAGTGAAGTGTGAGAGGACACACTCTGTGTCAGACACCATCG 1107

QY 508 AGCGATCTGGCCCTTACCGATGTCGATTTGCTGGACCTGAGTGGCTGCTAGACCGT 567
Db 1108 AGCGATCTGGCCCTTACCGATGTCGATTTGCTGGACCTGAGTGGCTGCTAGACCGT 1167
QY 568 CCATCTGTCAACACAGCCCAATGGT 593
Db 1168 CAGCCTCTCCCTATGCGCCACTGTGT 1193

RESULT 13
CD613789/c
LOCUS 56022208J1 FLP Homo sapiens cDNA, mRNA linear EST 12-JAN-2004
DEFINITION 56022208J1 FLP Homo sapiens cDNA, mRNA sequence.
ACCESSION CD613789
VERSION CD613789.1 GI:40262053
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 548)
AUTHORS Fu, G.K., Wang, J.T., Yang, J., Au-Young, J. and Stuve, L.L.
TITLE Circular rapid amplification of cDNA ends for high-throughput
extension cloning of partial genes
JOURNAL Genomics 84 (1), 205-210 (2004)
PUBMED 15203218
COMMENT Contact: Fu GK
Incyte Genomics, Inc.
3160 Porter Dr., Palo Alto, CA 94304, USA
Tel: 6508454102
Email: gfu@incyte.com.
Location/Qualifiers

FEATURES
source
1..548
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="FLP"
/note="Vector: pDrive Cloning Vector"

ORIGIN
Query Match 13.0%; Score 231.2; DB 6; Length 548;
Best Local Similarity 79.5%; Pred. No. 1.1e-54;
Matches 307; Conservative 0; Mismatches 33; Indels 46; Gaps 1;

QY 254 CCCATAGCTGAGAAGTGTGATCATGCTGCTGGGACTTCCTATGTGTCGAGAAACG 313
Db 440 CCCATAGCTGAGAAGTGTGATCATGCTGCTGGGACTTCCTATGTGTCGAGAAACG 381
QY 314 -----TGGGAGAGGCGC 327
Db 380 TGGGAGAGCCCTTACCAAGCCGGATGATGGTAGATTGTACTTGCCTGGGAGAGGCGC 321
QY 328 GGACGCATCACTTGCACTTCTAGAAATAGATGCAACGATCAGGACACAAGGACATCTAT 387
Db 320 GGACGCATCACTTGCACTTCTAGAAATAGATGCAACGATCAGGACACAAGGACATCTAT 261
QY 388 AGAATTGGAGACACCTGGAGCAAGAGGATPAATCGAGGAAACCTGCTCCAGTCCATCTGC 447
Db 260 AGAATTGGAGACACCTGGAGCAAGAGGATPAATCGAGGAAACCTGCTCCAGTCCATCTGC 201
QY 448 ACAGGCAACGGCCGAGGAGTGGAAAGTGTGAGAGGCACACCTCTGTGAGACCAATCG 507
Db 200 ACAGGCAACGGCCGAGGAGTGGAAAGTGTGAGAGGCACACCTCTGTGAGACCAATCG 141
QY 508 AGCGATCTGGCCCTTACCGATGTCGATTTGCTGGACCTGAGTGGCTGCTAGACCGT 567
Db 140 AGCGATCTGGCCCTTACCGATGTCGATTTGCTGGACCTGAGTGGCTGCTAGACCGT 81
QY 568 CCATCTGTCAACACAGCCCAATGGT 593
Db 80 CAGCCTCTCCCTATGCGCCACTGTGT 55

RESULT 14
CD613788
LOCUS 56022208H1 FLP Homo sapiens cDNA, mRNA linear EST 12-JAN-2004
DEFINITION 56022208H1 FLP Homo sapiens cDNA, mRNA sequence.
ACCESSION CD613788
VERSION CD613788.1 GI:40262052
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 551)
AUTHORS Fu, G.K., Wang, J.T., Yang, J., Au-Young, J. and Stuve, L.L.
TITLE Circular rapid amplification of cDNA ends for high-throughput
extension cloning of partial genes
JOURNAL Genomics 84 (1), 205-210 (2004)
PUBMED 15203218
COMMENT Contact: Fu GK
Incyte Genomics, Inc.
3160 Porter Dr., Palo Alto, CA 94304, USA
Tel: 6508454102
Email: gfu@incyte.com.
Location/Qualifiers

1..551
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="FLP"
/note="Vector: pDrive Cloning Vector"

ORIGIN

Query Match 13.0%; Score 231.2; DB 6; Length 551;
Best Local Similarity 79.5%; Pred. No. 1.1e-54;
Matches 307; Conservative 0; Mismatches 33; Indels 46; Gaps 1;
QY 254 CCCATAGCTGAGAAGTGTGATCATGCTGCTGGGACTTCCTATGTGTCGAGAAACG 313
Db 109 CCCATAGCTGAGAAGTGTGATCATGCTGCTGGGACTTCCTATGTGTCGAGAAACG 168
QY 314 -----TGGGAGAGGCGC 327
Db 169 TGGGAGAGCCCTTACCAAGCCGGATGATGGTAGATTGTACTTGCCTGGGAGAGGCGC 228
QY 328 GGACGCATCACTTGCACTTCTAGAAATAGATGCAACGATCAGGACACAAGGACATCTAT 387
Db 229 GGACGCATCACTTGCACTTCTAGAAATAGATGCAACGATCAGGACACAAGGACATCTAT 288
QY 388 AGAATTGGAGACACCTGGAGCAAGAGGATPAATCGAGGAAACCTGCTCCAGTCCATCTGC 447
Db 289 AGAATTGGAGACACCTGGAGCAAGAGGATPAATCGAGGAAACCTGCTCCAGTCCATCTGC 348
QY 448 ACAGGCAACGGCCGAGGAGTGGAAAGTGTGAGAGGCACACCTCTGTGAGACCAATCG 507
Db 349 ACAGGCAACGGCCGAGGAGTGGAAAGTGTGAGAGGCACACCTCTGTGAGACCAATCG 408
QY 508 AGCGATCTGGCCCTTACCGATGTCGATTTGCTGGACCTGAGTGGCTGCTAGACCGT 567
Db 409 AGCGATCTGGCCCTTACCGATGTCGATTTGCTGGACCTGAGTGGCTGCTAGACCGT 468
QY 568 CCATCTGTCAACACAGCCCAATGGT 593
Db 469 CAGCCTCTCCCTATGCGCCACTGTGT 494

RESULT 15
AUI40971
LOCUS AUI40971 PLACE4 Homo sapiens cDNA clone PLACE4000583 5', mRNA linear EST 05-AUG-2002
DEFINITION AUI40971 PLACE4 Homo sapiens cDNA clone PLACE4000583 5', mRNA sequence.
ACCESSION AUI40971
VERSION AUI40971.1 GI:11002492

KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 861)
AUTHORS Ota,T., Suzuki,Y., Saito,K., Ishii,S., Yamamoto,J., Sugiyama,T.,
Nishikawa,T., Nakamura,Y., Sugano,S., Masuho,Y. and Isoqai,T.
TITLE HRI human cDNA project (Ota,T., Suzuki,Y., Saito,K., Ishii,S.,
Yamamoto,J., Sugiyama,T., Nishikawa,T., Nakamura,Y., Sugano,S.,
Masuho,Y., Isoqai,T.)
JOURNAL Unpublished (2000)
COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
FEATURES
source Location/Qualifiers
1..861
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACB4000583"
/tissue_type="placenta"
/clone_lib="PLACE4"
/note="Vector: pME18SFL3"

ORIGIN

Query Match 13.0%; Score 231.2; DB 1; Length 861;
Best Local Similarity 79.5%; Pred. No. 1.3e-54;
Matches 307; Conservative 0; Mismatches 33; Indels 46; Gaps 1;
QY 254 CCCATAGCTGAGAGTGTGTTGATCATGCTGCTGGGACTTCCTATGTGGTGGGAGAAACG 313
DB 158 CCCATAGCTGAGAGTGTGTTGATCATGCTGCTGGGACTTCCTATGTGGTGGGAGAAACG 217
QY 314 -----TGGGAGAGGCGAGC 327
DB 218 TGGGAGAGGCCCTACCAAGCTGGATGATGTTAGATTGCTTGGGAGAGGCGAGC 277
QY 328 GGAACGATCATTGCACTTCTAGAAATAGATGCAACGATCAGGACACAAAGGACATCCTAT 387
DB 278 GGACGCATCATTGCACTTCTAGAAATAGATGCAACGATCAGGACACAAAGGACATCCTAT 337
QY 388 AGAATTGGAGACACCTGGAGCAAGAGGATAATCGAGGAAACCTGTCTCCAGTGCATCTGC 447
DB 338 AGAATTGGAGACACCTGGAGCAAGAGGATAATCGAGGAAACCTGTCTCCAGTGCATCTGC 397
QY 448 ACAGGCAACGGCCGAGGAGTGGAGTGGAGGCGACACCTGTGCGAGACACATCG 507
DB 398 ACAGGCAACGGCCGAGGAGTGGAGTGGAGGCGACACCTGTGCGAGACACATCG 457
QY 508 AGCGGATCTGGCCCTTACCGATGTTGCTGGACCTGAGTGGCTGTAGACCGT 567
DB 458 AGCGGATCTGGCCCTTACCGATGTTGCTGGACCTGAGTGGCTGTAGACCGT 517
QY 568 CCATCTGTCAACAAACAGCCAAATTGGT 593
DB 518 CAGCCTCTCCATATGGCCACTGTGT 543

Search completed: February 1, 2006, 12:22:51
Job time : 7324.56 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 28, 2006, 01:47:34 ; Search time 123.96 Seconds
(without alignments)

12632.698 Million cell updates/sec

Title: US-09-940-235-11

Perfect score: 3134

Sequence: 1 tcgcttcacgttcgtcgcg.....ataacctaacgacaaataa 1782

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 4886326

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_n2p model -DEV=xlp
-Q/cn2_1/USFTO.spool_p/US09940235/runat 27012006_144217_27548/app_query.fasta_1.7708
-DB=A_Geneseq -QFMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09940235 @CN1_1_846 @runat 27012006_144217_27548 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

A_Geneseq_21:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2002s:*
6: Geneseq2003as:*
7: Geneseq2003bs:*
8: Geneseq2004s:*
9: Geneseq2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2168	69.2	795	2 AAY24797	Aay24797 Streptoki
2	2165	69.1	483	2 AAR12885	Aar12885 Factor xa
3	2158	68.9	414	2 AAR10194	Aar10194 Streptoki
4	2158	68.9	414	2 AAR63120	Aar63120 Streptoki
5	2158	68.9	414	2 AAY24794	Aay24794 Streptoco
6	2158	68.9	414	2 AAW94664	Aaw94664 Streptoco
7	2158	68.9	414	2 AAY01556	Aay01556 Native st
8	2158	68.9	414	3 AAY90282	Aay90282 S. equisi
9	2157	68.8	531	2 AAR11829	Aar11829 FB-FB-SK

10	2154	68.7	414	2 AAW94665	Aaw94665 Streptoco
11	2153	68.7	413	8 ADL92189	Adl92189 Streptoki
12	2150.5	68.6	1194	2 AAW21726	Aaw21726 Streptoki
13	2143	68.4	499	2 AAR12891	Aar12891 Streptoki
14	2140.5	68.3	813	2 AAW21728	Aaw21728 Wild type
15	2132	68.0	414	2 AAW86143	Aaw86143 Streptoki
16	2132	68.0	414	3 AAB01295	Aab01295 Wild type
17	2132	68.0	415	9 AEB93425	Aeb93425 Streptoco
18	2132	68.0	440	2 AAR12889	Aar12889 Streptoki
19	2132	68.0	483	2 AAR12522	Aar12522 Factor xa
20	2132	68.0	859	2 AAR12893	Aar12893 OmpAL str
21	2128	67.9	414	2 AAR20202	Aar20202 S.equisim
22	2121	67.7	414	2 AAW86144	Aaw86144 De-immuni
23	2121	67.7	414	3 AAB01296	Aab01296 Altered s
24	2120.5	67.7	415	3 AAY50870	Aay50870 Streptoco
25	2120.5	67.7	415	3 AAY99593	Aay99593 Streptoco
26	2120.5	67.7	1194	2 AAW21724	Aaw21724 Modified
27	2116	67.5	440	2 AAR06377	Aar06377 Streptoki
28	2114.5	67.5	1194	2 AAY25020	Aay25020 Modified
29	2090.5	66.7	413	2 AAB80012	Aab80012 Streptoki
30	2090.5	66.7	413	5 ABB80012	Abb80012 Streptoki
31	2090.5	66.7	413	6 AEG74199	Aeg74199 Represent
32	2072	66.1	414	4 AAB74940	Aab74940 Mutant st
33	2071	66.1	800	2 AAW21723	Aaw21723 Plasminog
34	2071	66.1	1181	2 AAW21727	Aaw21727 Streptoki
35	2061	65.8	401	3 AAY84004	Aay84004 Amino aci
36	2061	65.8	413	3 AAY84005	Aay84005 Amino aci
37	2022	64.5	440	1 AAP50620	Aap50620 Sequence
38	2005.5	64.0	391	2 AAR10199	Aar10199 Streptoki
39	1938	61.8	440	2 AAR06378	Aar06378 Streptoki
40	1937	61.8	440	5 ABP29561	Abp29561 Streptoco
41	1937	61.8	440	8 ADR83961	Adr83961 S. pyogen
42	1924	61.4	372	2 AAR10200	Aar10200 Streptoki
43	1916	61.1	372	2 AAR10197	Aar10197 Streptoki
44	1907.5	60.9	371	2 AAR10195	Aar10195 Streptoki
45	1907	60.8	374	2 AAR10198	Aar10198 Streptoki

ALIGNMENTS

RESULT 1

AAY24797

ID AAY24797 standard; protein; 795 AA.

AC AAY24797;

XX XX 26-AUG-1999 (first entry)

XX Streptokinase and maltose binding protein fusion protein.

XX Streptococcus; streptokinase; fibrin-dependent plasminogen activator;
nSK; rSK; bacterial; blood clot; thrombotic condition;
myocardial infarction; venous thrombosis; pulmonary embolism;
cerebral thrombosis; graft thrombosis; arterial thrombosis.

XX OS Streptococcus dysgalactiae subsp. equisimilis.
OS Synthetic.

XX WO9931247-A1.

XX PD 24-JUN-1999.

XX PF 15-DEC-1998; 98WO-US026694.

XX PR 15-DEC-1997; 97US-0069497P.

XX PA (HARD) HARVARD COLLEGE.

XX PI Reed GL;

XX DR WPI; 1999-395183/33.

XX DR N-PSDB; AAX80497.

PT N-terminally deleted streptokinase.
 XX Example; Page 48-51; 73pp; English.
 XX
 CC The present invention describes an isolated bacterial protein that
 CC induces fibrin-dependent plasminogen activation in a pharmaceutical
 CC composition for dissolving blood clots. Also described are: (1) a
 CC composition comprising an isolated modified streptokinase, the
 CC modification being removal of amino acid residues in the amino terminus;
 CC (2) a method for dissolving a blood clot in a subject, comprising
 CC administering to the subject a fibrin-dependent streptokinase protein; a
 CC nucleic acid (1) encoding a modified bacterial streptokinase; (3) an
 CC expression vector comprising (1); and (4) a host cell transformed with
 CC the expression vector of (3). The pharmaceutical composition comprising a
 CC bacterial fibrin-dependent plasminogen activator is useful for dissolving
 CC blood clots in patients with a thrombotic condition, e.g. myocardial
 CC infarction, venous thrombosis, pulmonary embolism, cerebral thrombosis,
 CC graft thrombosis and arterial thrombosis. The modified streptokinase can
 CC also be used in non-human mammals. Streptokinase activation of
 CC plasminogen is at least 10-fold, preferably 100-fold greater in the
 CC presence of fibrin than in the absence of fibrin. The modified
 CC streptokinase has at least one amino acid substitution that inactivates a
 CC substrate site for proteolytic cleavage. This reduces the rate of
 CC degradation of the streptokinase at least two-fold. The present sequence
 CC represents a streptokinase and maltose binding protein fusion protein
 CC from an example of the present invention
 XX
 SQ Sequence 795 AA;

Alignment Scores:
 Pred. No.: 1,45e-193 Length: 795
 Score: 2168.00 Matches: 421
 Percent Similarity: 97.91% Conservative: 1
 Best Local Similarity: 97.68% Mismatches: 5
 Query Match: 69.18% Indels: 4
 DB: 2 Gaps: 2

US-09-940-235-11 (1-1782) x AAY24797 (1-795)

QY 496 CAGACCATCGAGC-----GGATCTGGCCCTTCCCGATGTTTGGTATTGCTGA 546
 DB 366 GlnThrAsnSerSerValProGlyArgGlySerIle---GluGlyArgIleAlaGly 384
 QY 547 CCTGAGTGGCTGAGCGCTCCATCTGTCACACAGCCCAATTTGGTGTCTAGCGTTGCT 606
 DB 385 ProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuValSerValAla 404
 QY 607 GGTACTGTTGAGGGACGAATCAAGACATTAGTCTTAAATTTTGGAAATCGATCTAAACA 666
 DB 405 GlyThrValGluGlyThrAsnGlnAspIleSerLeuLeuPhePheGluIleAspLeuThr 424
 QY 667 TCACGACCTGCTCATGGAGGAAGACAGAGCGCTTAAGTCCAAATCAAAACCATTT 726
 DB 425 SerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSerLysProPhe 444
 QY 727 GCTACTGATAGTGGCGGATGTCACATAAATTGAGAAAGCTGACTTACTAAAGGCTATT 786
 DB 445 AlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeuLysAlaIle 464
 QY 787 CAAGAACCAATTGATCGCTAACCTCCACAGTAACTGAGAGCTACTTTGAGGTCAATTGATTTT 846
 DB 465 GlnGluGlnLeuLeuAlaAsnValHisSerAsnAspAspTyrPheGluValIleAspPhe 484
 QY 847 GCAAGCGATGCAACCATTAATCTGATCGAAACGGCAAGGCTTACTTGTCTGCAAGAATGTT 906
 DB 485 AlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAspLysAspGly 504
 QY 907 TCGGTAACTTGGCCGACCACTGTCAGAGATTTTTCCTAAGCGGACATGTCGCCGTT 966
 DB 505 SerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHisValArgVal 524
 QY 967 AGACCATATAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGCTGGAATATCT 1026

DB 525 ArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspValGluTyrThr 544
 QY 1027 GTACAGTTTACTCCCTTAAACCCCTGATGACGATTTCAGACAGCTCTCAAGATACTAAG 1086
 DB 545 ValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLysAspThrLys 564
 QY 1087 CTATTGAAAAACACTAGCTATCGGTGACACCATCACATCTCTCAAGAATTACTAGCTCAAGCA 1146
 DB 565 LeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeuAlaGlnAla 584
 QY 1147 CAAGCATTTTAAACAAAACCCAGGCTATACGATTTATGAACTGATGCTCTCAATC 1206
 DB 585 GlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAspSerIle 604
 QY 1207 GTCACTCATGACATGACATTTCCGTACGATTTTACCAATGATGATCAAGATTTACTTAC 1266
 DB 605 ValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGluPheThrTyr 624
 QY 1267 CGTGTAAAAAATCGGAACCAAGCTTATAGGATCAATAAAAAAATCTGCTCGAATGAAGAA 1326
 DB 625 ArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeuAsnGluGlu 644
 QY 1327 ATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGGGAAAGCCG 1386
 DB 645 IleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGlyGluLysPro 664
 QY 1387 TATGATCCCTTTGATCGAGTCACTTGAACCTGTTCCACCATCAATATACGTTGATGTCAT 1446
 DB 665 TyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrValAspValAsp 684
 QY 1447 ACCAACCAATTGCTAAAAAGTGAAGAGCTCTTAAACAGTAGCGAACCTAACTTAGACTTC 1506
 DB 685 ThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsnLeuAspPhe 704
 QY 1507 AGAGATTATACGATCTCTGTAAGGCTAAACTACTCTACAACTCTCGATGCTGTTTT 1566
 DB 705 ArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeuAspAlaPhe 724
 QY 1567 GGTATTATGACATTAACCTTAACCTGGAAGAGTAGAGATTAATCAGATGACCAACCGT 1626
 DB 725 GlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAspThrAsnArg 744
 QY 1627 ATCATAAACGCTTATATGGCAAGCGACCCGAGAGGAGAGATCTAGCTATCATTTAGCC 1686
 DB 745 IleIleThrValTyrMetGlyLysArgProGluGluArgGluValTyrSerTyrLeuArgTyrThr 784
 QY 1687 TATGATAAAGATCGTTTATACCGAAGAGAACGAGAGATTTTACAGTACCTCGCTTATACA 1746
 DB 765 TyrAspLysAspArgTyrThrGluGluArgGluValTyrSerTyrLeuArgTyrThr 784
 QY 1747 GGGACACCTATACCTGATTAACCTTAACCGACAAA 1779
 DB 785 GlyThrProIleProAspAsnProAsnAspLys 795
 RESULT 2
 AAR12885
 ID AAR12885 standard; protein; 483 AA.
 XX AAR12885;
 AC
 XX
 XX 25-MAR-2003 (revised)
 DT 17-SEP-1991 (first entry)
 DT
 XX
 DE Factor Xa-cleavable hirudin-IEGR-streptokinase.
 XX
 XX Fusion protein; blood clotting; coagulation; fibrinolysis;
 KW antithrombotic; thrombolysis.
 XX
 OS Synthetic.
 XX
 FH Key
 FT Protein
 FT Location/Qualifiers
 FT 1..65
 FT /label= hirudin HV-1
 FT

QY 1597 CTAGAGGATAATCAGCATGACACCAACCGTATCATACCGTTTATATGGCAAGCGACCC 1656
 Db ValGluAspAsnHisAspAspThrAsnArgilleThrValTyrMetGlyLysArgPro 442

QY 1657 GAGAGAGAGATGCTAGCTATCATTTAGCCTATGATAAGATCGTTATACCGAAGAGAA 1716
 Db GluGlyGluAsnAlaSerTyrHisLeuAlaTyrAspLysAspArgTyrThrGluGluGlu 462

QY 1717 CGAGAGGTTTACAGCTACCTGCTGTTATACAGGACACCTATACCTGATACCTACCGAC 1776
 Db ArgGluValTyrSerTyrLeuArgTyrThrGlyThrProIleProAspAsnProAsnAsp 482

QY 1777 AAA 1779
 Db 483 Lys 483

RESULT 3
 AAR10194
 ID AAR10194 standard; protein; 414 AA.
 AC AAR10194;
 XX 28-MAR-1991 (first entry)
 DT Streptokinase encoded by synthetic gene.
 XX Streptokinase; thrombolytic agent; myocardial infarction.
 KW Synthetic.
 OS EP407942-A.
 PN 16-JAN-1991.
 PD 11-JUL-1989; 89JP-00179432.
 XX 11-JUL-1989; 89JP-00179432.
 PR 27-NOV-1989; 89JP-00307957.
 PR 11-APR-1990; 90JP-00096830.
 XX (SAKA) OTSUKA PHARM FACTOR.
 PA Majima E, Ogino K, Ono K, Sakata Y, Uenoyama T;
 PI WPI; 1991-016179/03.
 XX N-PSDB; AAQ10230.
 DR Synthetic gene encoding streptokinase - scale, high purity prodn. of
 XX streptokinase used as a thrombolytic agent.
 PT Claim 1; Page 51; 76pp; English.
 FS Streptokinase and its derivatives can be produced in large quantities
 CC with high purity for use as thrombolytic agents in patients with lung
 CC thrombus or myocardial infarction. See also AAR10195-R10200
 XX Sequence 414 AA;
 SQ

Alignment Scores:
 Pred. No.: 9,55e-193 Length: 414
 Score: 2158.00 Matches: 414
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 68.86% Indels: 0
 DB: 2 Gaps: 0

US-09-940-235-11 (1-1782) x AAR10194 (1-414)

QY 538 ATTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCCAATTGGTGT 597
 Db 1 ileAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal 20

QY 598 AGCGTTGCTGCTACTGTTGAGGGAGCAATCAAGACATTAGTCTTAATATTTTGAATC 657
 Db 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 40

QY 658 GATCTAACATCAGACCTGCTCATGAGGAGAAAGACAGACGAGCTTAAGTCCGAATCA 717
 Db 41 AspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 60

QY 718 AAACCATTTCTACTGATAGTGGCGGATGTCATATAACTTGAGAAGCTGACTTACTA 777
 Db 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeu 80

QY 778 AAGCGTATTCAAGAACAAATTCGCTAACCGTCCACAGTAAACGACGACTACTTTGAGGTC 837
 Db 81 LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100

QY 838 ATTGATTTTCCAAGCGATGCAACATTACTGTATGCGAAACGGCAAGGCTCTCTTCTGTCAC 897
 Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120

QY 898 AAGATGCTTCGGTAACTTGGCGACCCCAACCTGTCCAAGAAATTTTGTAAAGCGGACAT 957
 Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140

QY 958 GTGCGGTTAGACCATATATAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 1017
 Db 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160

QY 1018 GAATATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTCAGACGAGTCTCAA 1077
 Db 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180

QY 1078 GATACTAAGCTATTGAACACTAGCTATCGGTGACACCATCATCATCTCAAGATTACTA 1137
 Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 200

QY 1138 GCTCAAGACAAAGCATTTTAAACAAACCAACCCAGGCTATACGATTATGAACGTGAC 1197
 Db 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 220

QY 1198 TCCTCAATCGTCTCATGACAATGACATTTTCGTACGATTTTACCAATGGATCAAGAG 1257
 Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240

QY 1258 TTTACTTACCGTGTAAATAATCGGAAACAGCTTATAGGATCAATAAAATCTGGTCTG 1317
 Db 241 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 260

QY 1318 AATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 1377
 Db 261 AsnGluGluLeuAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280

QY 1378 GAAAGCGGTATGATCCCTTTGATCGCAGTCACCTTGAACCTGTTCCACCATCAATAGCTT 1437
 Db 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300

QY 1438 GATGTCGATACCAACGAATTCCTAAAGTGAGGAGCTCTTAACAGCTAGCAACGTAAC 1497
 Db 301 AspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320

QY 1498 TTAGACTTCAGAGATTTATACGATCCTCGGTGATAAGGCTAAACTTACTCTACAACTC 1557
 Db 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340

QY 1558 GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAAAGTAGAGGATAATACGATGAC 1617
 Db 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360

QY 1618 ACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAGAGAGAGATGCTAGCTAT 1677
 Db 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 380

QY 1678 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAACGAGAGAGTTTACAGCTACCTG 1737

Db 381 HisLeuAlaTyrAspLysAspArgTyrThrGluGluArgGluValTyrSerTyrLeu 400
QY 1738 CGTTATACAGGACACCTATACCTGATAACCCCTAAACGACAAA 1779
Db 401 ArgTyrThrGlyThrProIleProAspAsnProAsnAspLys 414

RESULT 4

AAR63120
ID AAR63120 standard; protein; 414 AA.
XX AAR63120;
AC AAR63120;
XX 16-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 16-NOV-1994 (first entry)
XX Streptokinase.
XX Streptokinase; SK; Streptococcus equisimilis; plasminogen;
KW myocardial infarction.
XX Streptococcus dysgalactiae subsp. equisimilis.

Key	Location/Qualifiers
FT Region	1..352
FT Region	/note= "claim 3, see CC"
FT Region	14..414
FT Region	/note= "claim 1, see CC"
FT Region	120..352
FT Region	/note= "claim 3, see CC"
FT Region	244..414
FT Region	/note= "claim 3, see CC"
FT Region	244..352
FT Region	/note= "claim 2, see CC"

W09407992-A1.

14-APR-1994.

05-OCT-1993; 93WO-US009502.

05-OCT-1992; 92US-00956692.

29-SEP-1993; 93US-00128299.

(GEO) GEN HOSPITAL CORP.

(HARD) HARVARD COLLEGE.

Reed GL;

WPI; 1994-135561/16.

DNA encoding a polypeptide which binds to plasminogen and corresponds to

region of streptokinase - useful to detect plasminogen in a sample and to

treat myocardial infarction.

Disclosure; Page 40-41; 62pp; English.

Nucleic acid comprising a sequence encoding amino acids 14-414 of

streptokinase, which binds to plasminogen and does not have a sequence

comprising amino acids 60-414 is new. The polypeptide pref. comprises

amino acids 244-352, 1-352, 120-352 or 244-414. (Updated on 25-MAR-2003

to correct PN field.) (Updated on 16-OCT-2003 to standardise OS field)

Sequence 414 AA;

Alignment Scores:

Pred. No.: 9,55e-193 Length: 414

Score: 2158.00 Matches: 414

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 68.86% Indels: 0

DB: 2 Gaps: 0

US-09-940-235-11 (1-1782) x AARE3120 (1-414)

QY 538 ATTCGTGACCTGAGTGGCTCTAGACCGTCCATCTGTCAACACAGCCAAATGGTGGTT 597
Db 1 IleAlaGlyProGluTrpLeuAspArgProSerValAsnAsnSerGlnLeuVal 20

QY 598 AGCGTTGCTGTACTGTTGAGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAATC 657
Db 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPheGluLeu 40

QY 658 GATCTAATCATCAGCACTGCTCATGAGGAAAGACAGAGCAAGCGCTTAAGTCCAAATCA 717
Db 41 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 60

QY 718 AAACCATTTGCTACTGATAGTGGCGGATGTCACATAAATCTGAGAAGCTGACTTACTA 777
Db 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 80

QY 778 AAGCTATTCAAGAAACAATTGATCGCTAAACGTCCACAGTACAGCAGCTACTTTGAGTC 837
Db 81 LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100

QY 838 ATTCATTTTGCAGCGATGCAACCATTAATCTGATCGAAACGCCAAGGCTACTTTCTGAC 897
Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120

QY 898 AAAGATGGTTCGTAACCTTCCGACCCCACTCTCCAGAAATTTTGGCTAAGCGGACAT 957
Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140

QY 958 GTGGCGTTAGACCATATAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 1017
Db 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160

QY 1018 GAATATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTACAGCAGGCTCTCAA 1077
Db 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuLys 180

QY 1078 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTA 1137
Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 200

QY 1138 GCTCAAGCACAAAGCATTTTAAACAAAACCAACCCAGGCTATACGATTTATGACGTCAC 1197
Db 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 220

QY 1198 TCCTCAATCGTCACTCATGACAATGACATTTTCGTACGATTTTACCAATGATCAAGAG 1257
Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240

QY 1258 TTTACTTACCGTGTAAAAATCGGGAACAAAGCTTATAGGATCAATAAAAAATCTGGTCTG 1317
Db 241 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 260

QY 1318 AATGAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1377
Db 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280

QY 1378 GAAAGCCGCTATGATCCCTTTTGTATCGCAGTCACCTTGAACCTGTTCCACCATCAATAGCT 1437
Db 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300

QY 1438 GATGTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCAGCACTAAC 1497
Db 301 AspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320

QY 1498 TTAGACTTCAGAGATTTATAGATCTCTCGTGATAGGCTAAACTACTCTACAAACATCTC 1557
Db 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340

QY 1558 GATGCTTTTGGTATTATGGACTACTACCTTAACCTGAAAGTAGAGGATATACGATGAC 1617

Db 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360

QY 1618 ACCAACCGTATCATACCGCTTTATATGGCGAAGCGACCCGAGAGAGAGATGCTAGCTAT 1677

Db 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 380

QY 1678 CATTTAGCCTATGATAAGATCGTTATATACCGAAGAGAAACGAGAAGTTTACAGCTACCTG 1737

Db 381 HisLeuAlaTyrAspLysAspArgTyrThrGluGluGluArgGluValTyrSerTyrLeu 400

QY 1738 CGTTATACAGGACACCTATACCTGATTAACCTTAACGACAAA 1779

Db 401 ArgTyrThrGlyThrProIleProAspAsnProAsnAspLys 414

RESULT 5

ID AAY24794 standard; protein; 414 AA.

XX AAY24794;

AC XX

DT 17-OCT-2003 (revised)

DT 26-AUG-1999 (first entry)

XX Streptococcus equisimilis native streptokinase.

DE Streptococcus; streptokinase; fibrin-dependent plasminogen activator;

KW nSK; rSK; bacterial; blood clot; thrombotic condition;

KW myocardial infarction; venous thrombosis; pulmonary embolism;

KW cerebral thrombosis; graft thrombosis; arterial thrombosis.

XX Streptococcus dysgalactiae subsp. equisimilis.

OS WO9931247-A1.

XX 24-JUN-1999.

PD 15-DEC-1998; 98WO-US026694.

PF 15-DEC-1997; 97US-0069497P.

PR (HARD) HARVARD COLLEGE.

XX Reed Glu;

XX WPI; 1999-395183/33.

DR N-PSDB; AAX80492.

XX N-terminally deleted streptokinase.

PS Claim 30; Page 60-61; 73pp; English.

CC The present invention describes an isolated bacterial protein that induces fibrin-dependent plasminogen activation in a pharmaceutical composition for dissolving blood clots. Also described are: (1) a composition comprising an isolated modified streptokinase, the modification being removal of amino acid residues in the amino terminus; (2) a method for dissolving a blood clot in a subject, comprising administering to the subject a fibrin-dependent streptokinase protein; a nucleic acid (I) encoding a modified bacterial streptokinase; (3) an expression vector comprising (I); and (4) a host cell transformed with the expression vector of (3). The pharmaceutical composition comprising a bacterial fibrin-dependent plasminogen activator is useful for dissolving blood clots in patients with a thrombotic condition, e.g. myocardial infarction, venous thrombosis, pulmonary embolism, cerebral thrombosis, graft thrombosis and arterial thrombosis. The modified streptokinase can also be used in non-human mammals. Streptokinase activation of plasminogen is at least 10-fold, preferably 100-fold greater in the presence of fibrin than in the absence of fibrin. The modified streptokinase has at least one amino acid substitution that inactivates a substrate site for proteolytic cleavage. This reduces the rate of degradation of the streptokinase at least two-fold. The present sequence represents native streptokinase (nSK). (Updated on 17-OCT-2003 to standardise OS field)

XX SQ Sequence 414 AA;

Alignment Scores: 9.55e-193 Length: 414

Score: 2158.00 Matches: 414

Percent Similarity: 100.00% Conservativeness: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 68.86% Indels: 0

DB: 2 Gaps: 0

US-09-940-235-11 (1-1782) x AAY24794 (1-414)

QY 538 ATTGCTGGACCTCGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCCAATGTTGTT 597

Db 1 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20

QY 598 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAATATTTTGAATC 657

Db 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 40

QY 658 GATCTAACATCAGACCTGCTCATGGAGGAAACAGACAGCAAGGCTTAAAGTCCAAATCA 717

Db 41 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 60

QY 718 AAACCAATTTGCTACTGATAGTGGCGGATGCTCATATAAATCTTGAGAAAGCTGCTACTA 777

Db 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeu 80

QY 778 AAGGCTATTCAAGAACAAATTTGATCGCTTAACGTCACAGTAACGACCACTACTTTGAGGTC 837

Db 81 LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100

QY 838 ATTGATTTTCAAGCGATGCAACATTACTGATCGAAACCGGAGGCTACTTTGCTGAC 897

Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120

QY 898 AAAGATGTTTCGGTAACTTGGCGACCCACCTGTCACGAATTTTGTCTAAGCGGACAT 957

Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140

QY 958 GTGCGGTTAGACCATATAAAGAAAAACCAATACAAACCAACGCAATCTGTTGATGTG 1017

Db 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160

QY 1018 GAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTCAGACAGGCTCTCAA 1077

Db 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180

QY 1078 GATACTAAGCTATTGAACACACTAGCTATCGGTGACACCATCATCTCAAGAACTACTA 1137

Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 200

QY 1138 GCTCAAGCACAAAGCATTTTAAACCAAAACCCAGGCTATACGATTATGACCGTAC 1197

Db 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 220

QY 1198 TCCTCAATCGTCTCATGCAATGATCTTTCCGTCAGATTTTACCAATGGATCAAGAG 1257

Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240

QY 1258 TTTACTACCGTGTAAAAATTCGGGAACAAGCTTATAGGATCAATAAAAAATCTGCTCTG 1317

Db 241 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 260

QY 1318 AATGAAGAATAAACCAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1377

Db 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280

QY 1378 GAAAGCCGTATGATCCCTTTTGATCGCAGTCACTTGAAATCTGTTCAACATCAATAGCTT 1437

Db 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300

QY 1438 GATGTCGATACCAACGAATTGCTAAAGAGAGAGCTTTAAACGCTACGAAACGTAC 1497
 DB 301 AspValAspThrAsnGluLeuLeuLeuSerGluLeuLeuThrAlaSerGluArgAsn 320
 QY 1498 TTAGACTTCAGAGATTATACGATCCTCGGATTAAGCTAACTACTCTACACAAATCTC 1557
 DB 321 LeuAspPheArgSerLeuValProAlaGlyAlaLeuLeuValLeuValLeuValLeu 340
 QY 1558 GATGCTTTGGTGTATGAGCTATACCTTAACCTGAAAGAGATAGATATTCACGATAC 1617
 DB 341 AspAlaPheGlyIleLeuLeuLeuValThrLeuThrGlyValGluAspAsnHisAspAsp 360
 QY 1618 ACCAACCGTATCATTAACCGTTATATGAGGACGACCGACGAGAGAAATGCTAGCTAT 1677
 DB 361 ThrAsnArgIleIleThrValIleMetGlyLeuArgProGluGluAsnAlaSerTyr 380
 QY 1678 CATTTAGCTTATGATTAAGATCGTTATACCGAAGAGAAAGAAAGATTACAGCTACCTG 1737
 DB 381 HisLeuAlaIleValAspValAspArgTyrThrGluGluGluValArgIleValTyrSerTyrLeu 400
 QY 1738 CGTTATACAGGAGACCTTATACCTGATTAACCTTAACGAGACAA 1779
 DB 401 ArgTyrThrGlyThrProIleProAspAsnProAsnAspLys 414
 RESULT 6
 AAW94664
 ID AAW94664 standard; protein; 414 AA.
 XX AAW94664;
 AC AAW94664;
 XX
 DT 17-OCT-2003 (revised)
 DT 04-MAY-1999 (first entry)
 DE Streptococcus equisimilis native streptokinase.
 XX Streptococcus equisimilis
 XX Streptococcus equisimilis H46A; streptokinase; mutant; fibrinolytic;
 KM plasma clot; hydrolysis; haemolytic Streptococcus; plasminogen; plasmin;
 KM serine protease; fibrin; blood clot; thrombolytic;
 KM vascular thromboembolytic symptom; acute myocardial infarction;
 KM fibrinolysis; resistance.
 OS Streptococcus dysgalactiae subsp. equisimilis.
 XX
 FH Key location/Qualifiers
 FT Misc-difference 14 /note="encoded by ACC"
 FT
 PN US5876999-A.
 XX
 PD 02-MAR-1999.
 XX
 PF 06-DEC-1995; 95US-00568393.
 XX
 PR 06-DEC-1995; 95US-00568393.
 XX
 PA (NASC-) NAT SCI COUNCIL.
 XX
 PI Wu H;
 XX
 DR MPI; 1999-189643/16.
 DR N-PSDB; AAX16632.
 XX
 PT Mutant streptokinase polypeptide - useful as plasmin-resistant
 PT thrombolytic agent.
 PS Claim 1; Col 7-10; 17pp; English.
 XX
 CC The present invention describes a mutant streptokinase (SK) polypeptide
 CC in which at least one of the amino acids in the Pro58-Lys59-Ser60-Lys61
 CC segment of the corresponding native SK is replaced by another amino acid.
 CC The present sequence represents native SK. SK is a secretory protein of
 CC haemolytic Streptococcus able to activate human plasminogen (HP1g) to
 CC plasmin (HP1m), which is a serine protease able to catalyse the

CC hydrolysis of fibrin in blood clots. The SK is useful as a thrombolytic
 CC agent in the treatment of vascular thromboembolytic symptoms such as
 CC acute myocardial infarction. Compared with wild-type SK, the K59E mutant
 CC is more resistant to degradation by human plasmin and is more effective
 CC both in acting as a fibrinolytic agent and in activating human plasminogen.
 CC (Updated on 17-OCT-2003 to standardise OS field)
 XX
 SQ Sequence 414 AA;
 Alignment Scores:
 Pred. NO.: 9.55e-193 Length: 414
 Score: 2158.00 Matches: 414
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 68.86% Indels: 0
 DB: 2 Gaps: 0
 US-09-940-235-11 (1-1782) x AAW94664 (1-414)
 QY 538 ATTGCTGAGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACGCAATTGGTGT 597
 DB 1 IleAlaGlyProGluIlePleuLeuAspArgProSerValAsnSerGluLeuVal 20
 QY 598 AGCGTTGCTGCTGCTGCTGTTGAGGGGAGCAATCAACATTAAGCTTTTGAATC 657
 DB 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuValPhePheGluIle 40
 QY 658 GATCTAACATCAGACACCTGCTCATGAGAGAAAGACAGAGCAAGGCTTAAGTCAAAATCA 717
 DB 41 AspleuThrSerArgProAlaHisGlyGlyThrGluGluGluLeuSerProLysSer 60
 QY 718 AAACCATTTGCTACTGATAGTGGCGGATGTCACATTAACCTTGAGAAAGCTGACTTACTA 777
 DB 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluValAlaAspLeuLeu 80
 QY 778 AAGGCTATTCAGAAACATTAATGATGCTTACGTCACAGTAAAGCACTACTTGAAGTGC 837
 DB 81 LysAlaIleGluGluGluGluLeuIleAlaAsnValHisSerAsnAspArgTyrPheGluVal 100
 QY 838 ATTGATTTTTCAGACGATGCAACCATTAAGTATGATGAAACGCGCAAGCTTACTTGTGTCAC 897
 DB 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyValTyrPheAlaAsp 120
 QY 898 AAGATGCTTGGTGGTAACTTGGCGACCCAACTGTCCAAGATTTTGTAAAGCGAACAT 957
 DB 121 LysAspGlySerValThrLeuProThrGluProValGluGluPheLeuSerGlyHis 140
 QY 958 GTGGCGGTTAGACCATTAAGAAACCAATCAACCAAGCAAGCAAGTGTGAAGTGTG 1017
 DB 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
 QY 1018 GAATATACGTATGACGTTACTCCCTTAACCTGATGACGATTTCAACAGGCTGTCAAA 1077
 DB 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
 QY 1078 GATACTAAGCTATTAAGAAACATGACTATGCGTGACACCATCATCATCAAGATTACTA 1137
 DB 181 AspThrLysLeuLeuValThrLeuAlaIleGlyAspThrIleThrSerGluLeuLeu 200
 QY 1138 GCTCAACGACAAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTATGAACGTGAC 1197
 DB 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 220
 QY 1198 TCCGATCGTCACTCATGACATGACATTTTCGCTGATTTTCAACAGATCAAGAG 1257
 DB 221 SerSerIleValIleThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
 QY 1258 TTTACTTACCGTGTAAATCGGAAACAAGCTTATAGATCAATAAATATGCTGTCTG 1317
 DB 241 PheThrTyrArgValLysAsnArgGluGlnAlaIleTyrArgIleAsnLysSerGlyLeu 260
 QY 1318 AATGAAGAAATTAACACACTGACCTGATCTTGAGAAATATTAAGCTTTAAAGAGG 1377

Db 281 GLuLySProTYrAsPProPheAsPAsGSeHisLeuLySLeuPheThrIleLySTyrVal 300
QY 1438 GATGTCGATACCAACCAATGCTTAAAGAGAGAGCTTTACAGCTAGCAAGCTAAC 1497
Db 301 AsPValAsPThrAsnIleuLeuLySLeuGluGlnLeuThrAlaSerGluArgAsn 320
QY 1498 TTAGACTTCAGAGATTATTCGATCTCTGATTAAGGCTAAACTACTCTTACAAATCTC 1557
Db 321 LeuAsPPhaArgSPLeuTYrAsPProArgAsPValAlaLySLeuLeuTYrAsnAsnLeu 340
QY 1558 GATGCTTTGGTATTATGACTTACCTTACCTGGAAGAGATATTCAGATGAC 1617
Db 341 AsPAlaPheGlyIleuLeuAsPTrpIleuThrGlyValGluAsPAsnHisAsPAsp 360
QY 1618 ACCAACCGTATCATACCGTTTATGAGGAGAGACCGGAGAGAGAGAAATGCTAGCTAT 1677
Db 361 ThrAsnArgIleIleThrValTYrMetGlyLyAsnArgProGluGluValAlaSerTYr 380
QY 1678 CATTAGCTTATGATTAAGATCGTTATACCGAAGAGAGAGAGAGTTTACAGCTACTG 1737
Db 381 HisLeuAlaIleTYrAsPValAsPArgTYrThrGluGluGluValTYrSerTYrLeu 400
QY 1738 CGTTATACGAGGACACTTACTGATTAACCTTACAGCAAA 1779
Db 401 ArgTYrThrGlyThrProIleProAsPAsnProAsnAsPlys 414
RESULT 8
AAV90282
ID AAV90282 standard; protein; 414 AA.
XX AAV90282;
AC AAV90282;
XX 12-SEP-2003 (revised)
DT 13-OCT-2000 (first entry)
XX
DE S. equisimilis streptokinase.
XX Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
KM plasminogen; human; fibrinolytic; thrombolytic therapy;
KM cardiovascular disorder; fibrinectin.
XX
OS Streptococcus dysgalactiae subsp. equisimilis.
XX
PN EP1024192-A2.
PD 02-AUG-2000.
XX
PF 23-DEC-1999; 99EP-00310541.
PR 24-DEC-1998; 98IN-DE003825.
PA (COUL) CSIR COUNCIL SCI IND RBS.
PI Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V,
PI Yadav M;
XX
DR WPI: 2000-516032/47.
DR N-PSDB; AAA37633.
XX
PT Hybrid streptokinase-fibrin binding domain polypeptides useful for
PT thrombolytic therapy comprises a streptokinase fused with fibrin binding
PT domains of human fibrinectin.
XX
XX Example 3; Fig 3; 58p; English.
XX This sequence represents the human Streptococcus equisimilis
XX streptokinase protein sequence. The invention relates to a hybrid
XX plasminogen activator (PA) comprises a polypeptide fusion between
XX streptokinase (SK), which are capable of plasminogen (PG) activation, and
XX fibrin binding regions of human fibrinectin, which are from fibrin
XX binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA possesses the
XX ability to bind with fibrin independently and also characteristically
XX retains a PG activation ability which becomes evident only after a

CC pronounced duration, or lag, after exposure of the PA to a suitable
CC animal or human PG. The hybrid streptokinase-fibrin binding domain
CC polypeptides are useful in thrombolytic therapy for various kinds of
CC cardiovascular disorders. The hybrids have enhanced fibrin selectivity as
CC well as kinetics of plasminogen activation that are distinct from that of
CC natural streptokinase in being characterised by a temporary delay, or lag
CC of several minutes in the natural rate of the catalytic conversion of
CC plasminogen to plasmin (i.e. delayed-action thrombolytic). The proteins
CC can bind tightly with fibrin in blood clots soon after introduction into
CC the vascular system without significantly activating the circulating
CC blood plasminogen to plasmin, thus aiding in the localisation of the
CC plasminogen activation process to the site of pathological thrombus. This
CC overcomes systemic plasminogen activation encountered during clinical use
CC of streptokinase. (Updated on 12-SEP-2003 to standardise OS field)
XX
SQ Sequence 414 AA;
Alignment Scores:
Pred. No.: 9.55e-193 Length: 414
Score: 2158.00 Matches: 414
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 68.86% Indels: 0
DB: 3 Gaps: 0
US-09-940-235-11 (1-1782) x AAV90282 (1-414)
QY 538 ATTGCTGACCTGAGTGCTGCTGACCGTCCATCTGTCACAAACAGCAATGGTGT 597
Db 1 IleAlaGlyProGluTrpLeuAsPArgProSerValaAsnSerGlnLeuValVal 20
QY 598 AGCGTGTGCTGACTGTTGAGGGAGCGAATCAAGACTTACTTAAATTTTGAATC 657
Db 21 SerValAlaGlyTrpValGluGlyThrAsnGlnAspIleSerLeuLySPheGluIle 40
QY 658 GATCTAACATCAGCACTGCTCATGAGAGAAAGACAGAGAGGCTTAAGTCCAAATCA 717
Db 41 AsPLeuThrSerArgProAlaHisGlyGlySerThrGluGlnGlyLeuSerProLySer 60
QY 718 AAGCATTTGCTACTGATAGTGCGGAGATGCATATAACTTGAAGAACTGACTTA 777
Db 61 LySProPheAlaTrpAsPArgSerGlyAlaMetSerHisLySLeuGluValaAsPLeu 80
QY 778 AAGCTATTTCAGAAACAATTGATGCTTACGTCACAGTAAAGCACTACTTTGAGGTC 837
Db 81 LySAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAsPArgPheGluVal 100
QY 838 ATTGATTTTGAAGAGATGCAACCATTAAGTATGAGAAACGGCAAGTCTACTTTGCTGAC 897
Db 101 IleSPheAlaSerAsPAlaThrIleThrAsPArgSerGlyValTYrPheAlaAsP 120
QY 898 AAGATGTTTGGTAACTTGCAGCCGAACTGTCAGAAATTTTGTAAAGCGGACAT 957
Db 121 LyAsPArgLySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
QY 958 GTGGCGCTTACACCATTAAGAAACCAATACAAACCAAGGAAATCTGTGATGTG 1017
Db 141 ValArgValArgProTYrLySLeuGlyPProIleGlnAsnGlnAlaLySerValaAsPVal 160
QY 1018 GAATATATCTGATACGTTACTCCCTTAAACCGTAGAGATTTTCAAGACGAGCTGAAA 1077
Db 161 GluTYrThrValGlnPheThrProLeuAsnProAsPAsPArgProGlyLeuLyS 180
QY 1078 GATACTAAGCTATTGAAACACTAGTATGCGGTGACACCATCATCATCAAGATTACTA 1137
Db 181 AsPThrLySLeuLeuLySThrLeuAlaIleGlyAsPThrIleThrSerGlnGluLeu 200
QY 1138 GCTCAAGCAACAAAGATTTTAAACAAACCAACCAAGGCTTATGATTTAAGACGTGAC 1197
Db 201 AlAGlnAlaGlnSerIleLeuAsnLyAsnHisPProGlyTYrThrIleTYrGluArgAsP 220
QY 1198 TCTCATCTCTCATGATGACATGACATTTTCCGTACGATTTTACCAAGATGACAGAG 1257

```

Db      221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleuProMetAspGlnGlu 240
QY      1258 TTTACTTCCCGTGTAAATAATGGGAGACAGCTTATAGATCAATAATAATGTCCTG 1317
Db      241 PheThrIYrAspValIleAsnAsnThrAspLeuIleSerGlnIleYrIYrValIleuIleuSlySgIY 260
QY      1318 AATGAGAAATTAACAACACAGCTGATCTCTGAGAAATATATACCTCTTAAATAAGG 1377
Db      261 AsnGlnGlnIleAsnAsnThrAspLeuIleSerGlnIleYrIYrValIleuIleuSlySgIY 280
QY      1378 GAAAACCGTATGATCTCTTATGCGACATCTGAAAATCTGTCCATCAATAACGTT 1437
Db      281 GlnIlySProIYrAspProPheAspArgSerHisIleuIleuPheThrIleuSlyYrVal 300
QY      1438 GATGTCGATACCAAGAAATTTGTTAAATAGTGAAGAGCTCTTAAACACTAGCGCAAGTAA 1497
Db      301 AspValAspThrAsnGlnIleuIleuSlySerGlnIleuIleuThrIleuIleuIleuIleuIleu 320
QY      1498 TTAGACTTCAGAGATTTATACGATCCTCGTATAGGCTTAACTACTTACCAACAACTTC 1557
Db      321 LeuAspPheAspIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 340
QY      1558 GATGCTTTGTATTATGACATACTTAACTGAGAAAGTAAAGAGATATACAGATGAC 1617
Db      341 AspIlePheGlyIleMetAspIYrThrIleuThrGlnIleuValGlnAspAsnHisAspAsp 360
QY      1618 ACCAACCCTATCATTAACGTTTATATGGGACAGCCGAGAGAGAGAAATGCTTACTT 1677
Db      361 ThrAsnArgIleIleThrValIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 380
QY      1678 CATTAGGCTATGATTAAGATGCTTATATACGAAAGAAAGAGAACTTTACGCTACCTG 1737
Db      381 HisIleuAlaIYrAspIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 400
QY      1738 CGTTATACAGGAGACCTATACCTGATTAACCTTAAACGACAAA 1779
Db      401 ArgIYrThrGlnIYrThrProIleProAspAsnProAspAspIleu 414

RESULT 9
AAR11829
ID      AAR11829 standard; protein, 531 AA.
XX      AAR11829;
AC      AAR11829;
DT      08-JUL-1991 (first entry)
DE      FB-FB-SK fusion conjugate.
XX      Fibrin-binding protein; fibrinolysis; intravascular thrombi; fibrinogen;
KM      streptokinase; fusion protein.
XX      Staphylococcus aureus.
OS      Staphylococcus aureus.
FH      Key      Location/Qualifiers
FT      Peptide      2..61
FT      Peptide      /label= FB monomer
FT      Peptide      62..119
FT      Peptide      /label= FB monomer
FT      Peptide      120..531
FT      Peptide      /label= streptokinase
XX      US5011686-A.
XX      30-APR-1991.
XX      15-NOV-1989; 89US-00437769.
XX      21-SEP-1987; 87US-00099242.
XX      (CREA-) CREATIVE BIOMOLEC.
XX      Pang RHL;
XX      PI
XX

```

```

DR      WPI, 1991-140198/19.
DR      N-PSDB; AAQ11651.
XX      PT      Imparting injectable fibrinolytic agent - with affinity for intravascular
XX      thrombus, by linking agent to fibrin binding domain.
XX      PS      Disclosure; Fig 5; 18pp; English.
XX      CC      The conjugate comprises an FB-FB dimer linked to streptokinase. The FB
XX      fragment has selective affinity for fibrin, low affinity for fibrinogen,
XX      CC      and minimal immunogenicity, imparting thrombus- targeting capability.
XX      CC      See also AAR11821 and AAR11828
XX      DB      Sequence 531 AA.
SQ      Alignment Scores:
Pred. No.: 1,31e-192 Length: 531
Score: 2157.00 Matches: 413
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 99.52% Mismatches: 0
Query Match: 68.83% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-11 (1-1782) x AAR11829 (1-531)
QY      535 CGTATTCCTGAGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCAATGGTT 594
Db      117 LysIleAlaGlyProGlnIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 136
QY      595 GTTAGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 654
Db      137 ValSerValAlaGlyThrValGlnIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 156
QY      655 ATGATCTTAAATCATCACAGCCTGCTCATGAGAAAGACAGAGCAAGGCTTAACTCAAAA 714
Db      157 IleAspIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 176
QY      715 TCAAAACCATTTGCTACTGATAGTGGCGGATGCTCATTAATCTTGAAGAAAGCTGACTTA 774
Db      177 SerIleProPheAlaThrAspSerGlyAlaMetSerHisIleuIleuIleuIleuIleuIleuIleu 196
QY      775 CTAAGGCTATTCAGAAACAAATGATCGCTACGTCAGAGTAAAGCACTACTTGGAG 834
Db      197 LeuIleuAlaIleuGlnIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 216
QY      835 GTCATTTGATTTTCAAGCGATGCAACCATTAATGATCGAAACGCGACGCTTACTTGGCT 894
Db      217 ValIleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlySValIYrPheAla 236
QY      895 GACAAAGATGCTTGGTAACTTGGCCGACCCAACTGTCCAAAGATTTTGTAAAGCGGA 954
Db      237 AspIleAspIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 256
QY      955 GATGCGGCTTGAACCATTAAGAAACCAATCAAAACCAAGGAAATCGTGTAT 1014
Db      257 HisValArgValArgProIYrIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 276
QY      1015 GTGGAATATATCTGTACAGTTTACTCTTAAACCTGATGACGATTTTCAAGCAAGCTTC 1074
Db      277 ValGlnIYrThrValGlnPheThrProIleuAsnProAspAspAspPheArgProGlnIleu 296
QY      1075 AAGATACCTAGCTATTGAAACACATGACTATGCTGACACCATCATCTCAAGAAATTA 1134
Db      297 LysAspThrIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 316
QY      1135 CTAGCTCAAGCAAGCAATTTTAAACAAACCAACCAAGCTTATAGATTTTGAAGCT 1194
Db      317 LeuAlaGlnAlaGlnIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 336
QY      1195 GACTCCCAATCGTCACTGACATGACAAATGATTTTCCGTACGATTTTACCAATGATCAA 1254
Db      337 AspSerSerIleValIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 356

```

QY 1255 GAGTTTACTTACCGGTGTTAAAAATCGGAAACAAGCTTATGATCAATAAAAATCTGGT 1314
Db 357 GIupheThrTyZrArVallySaenArGluGlnAlaTyZrArGlyLeuSlySerGly 376
QY 1315 CTGAATGAAGAATAAACAACAGCTGATCTCTGAGAAATTTTACGCTTAAAAAA 1374
Db 377 LeuAsnGluGlnLeuAsnAntArAspLeuIleSerGlnSlyTyTyValLeuLyS 396
QY 1375 GGGGAAAGACCGGTATGATCCCTTGAATCGAGTCATCTTGAACCTGTCACATCAATAC 1434
Db 397 GIyGlnLySProTyArSPrProPheArSPrSerHisLeuLySLeuPheThrIleLyTy 416
QY 1435 GTTGAATGATGATCAACAAGATGCTTAAAAAGTGAAGAGCTTAAACAGCTAAGCGAAGT 1494
Db 417 ValAspValAspThrAsnGlnLeuLeuLySserGlnGlnLeuLeuThrAlaSerGlnArG 436
QY 1495 AACTTGAAGCTGAGATTTTATACGATCCCTCGATGATGAGCTTAACTACTTACAAACAT 1554
Db 437 AsnLeuAspPheArSPrLeuTyArSPrArSPrAlaLySLeuLeuTyArAsnAsn 456
QY 1555 CTGGAATGCTTTGGTATTTATGAGCTATACCTTAACTGGAAGAGATGATTAATCACAGT 1614
Db 457 LeuAspAlaPheGlyIleMetAspTyThrLeuThrGlyLySValGlnAspAsnHisAsp 476
QY 1615 GACACCAACCGTATCATTAACCGTTTATATGAGCAAGCAAGCCGAAAGAGAAATGCTAGC 1674
Db 477 AspThrAsnArGlyIleThrValTyMetGlyLyAspArgProGlnGlnAlaSer 496
QY 1675 TATCATTTAGCTTATGATTAAGATGCTTATACCGAAGAGACGAGAGTTTACAGCTAC 1734
Db 497 TyRHsLeuAlaIleTyArSPrAspArGlyTyThrGlnGlnGlnArGlnValTySerTy 516
QY 1735 CTGGTATTAACGGGACACCTATACCTGATTAACCTTAACGACAA 1779
Db 517 LeuArGlyTyThrGlyThrProIleProAspAsnProAsnAspLyS 531

RESULT 10
ID AAM94665
AAM94665 standard; protein; 414 AA.
XX
AC AAM94665;
XX
DT 04-MAY-1999 (first entry)
XX
DE Streptococcus equisimilis mutant streptokinase K59E.
XX
KW Streptococcus equisimilis H46A; streptokinase; mutant; fibrinolytic;
KW plasma clot; hydrolysis; haemolytic Streptococcus; plasminogen; plasmin;
KW serine protease; fibrin; blood clot; thrombolytic;
KW vascular thromboembolytic symptom; acute myocardial infarction;
KW fibrinolysis; resistance.
XX
OS Streptococcus dysgalactiae subsp. equisimilis.
OS Synthetic.
XX
FT Key Location/Qualifiers
FT Misc-difference 14 /note= "encoded by ACC"
FT

US5876999-A.
XX
XX
XX 02-MAR-1999.
XX
XX 06-DEC-1995; 95US-00568393.
XX
XX 06-DEC-1995; 95US-00568393.
XX
XX (NASC-) NAT SCI COUNCIL.
XX
XX Mu H;
XX
XX WPI, 1999-189643/16.
XX
XX N-P8DB; AAX16633.

XX
PT Mutant streptokinase polypeptide - useful as plasmin-resistant
PT thrombolytic agent.
XX
XX
PS Claim 4; Col 11-14; 17pp; English.
XX
XX The present invention describes a mutant streptokinase (SK) polypeptide
CC in which at least one of the amino acids in the Pro58-Lys69-Ser60-Lys61
CC segment of the corresponding native SK is replaced by another amino acid.
CC The present sequence is mutant SK K59E. SK is a secretory protein of
CC haemolytic Streptococcus able to activate human plasminogen (HPIg) to
CC plasmin (HPIu), which is a serine protease able to catalyse the
CC hydrolysis of fibrin in blood clots. The SK is useful as a thrombolytic
CC agent in the treatment of vascular thromboembolytic symptoms such as
CC acute myocardial infarction. Compared with wild-type SK, the K59E mutant
CC is more resistant to degradation by human plasmin and is more effective
CC both in acting as a fibrinolytic agent and in activating human plasminogen
XX
SQ Sequence 414 AA;
XX
Alignment Scores:
Pred. No.: 2.26e-192 Length: 414
Score: 2154.00 Matches: 413
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.76% Mismatches: 0
Query Match: 68.73% Indels: 0
DB: 2 Gaps: 0
US-09-940-235-11 (1-1782) x AAM94665 (1-414)
QY 538 ATTGCTGACCTGAGTGGCTGCTGACCGTCCATCTGTCACAAACAGCCATTTGGTT 597
Db 1 IleAlaIleProGlnThrLeuLeuAspThrProSerValAsnSerGlnLeuValAl 20
QY 598 AGCGTTCGTGATCTGTTGAGGGGACCAATCAAGACATTAAGTTTAAATTTTGAATC 657
Db 21 SerValAlaGlyThrValGlnGlyThrAsnGlnAspIleSerLeuLySPhenGluIle 40
QY 658 GATCTAATCATCAGACCTGCTCATGAGGAAAGACAGCAAGCTTAAGTCCAAATCA 717
Db 41 AspLeuThrSerArProAlaHisGlyLyThrGlnGlnGlnLeuSerProGlnSer 60
QY 718 AAACATTTGCTACTGATAGTGGGCGCATGTCACATTAACCTTGAAGAGCTTACTCTA 777
Db 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLySLeuGlnValAspLeuLeu 80
QY 778 AAGGCTATTCAGAAACATTTGATGCTTACCTCAAGTACAGCACTACTTTAGAGTTC 837
Db 81 LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspArTyRPhenGlnVal 100
QY 838 ATTGATTTGCAAGCGATGCAACCATTAAGTATGATGAAAACGGCAAGTCTTACTTGCTGAC 897
Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlySValTyRPhenAlaAsp 120
QY 898 AAGAGTGTTCGGTAACTTGGCCGACCAACCTGTCACAAAGATTTTGTGAAGGACAT 957
Db 121 LysAspIleSerValThrLeuProThrGlnProValGlnGlnPheLeuLeuSerGlyHis 140
QY 958 GTGCGCGTTAGACCATTAAGAAACCAATACAAACCAAGCAAGCAATCGTTGATGATG 1017
Db 141 ValArgValArgProTyRlySglnLySProIleGlnAsnGlnAlaLySValAspVal 160
QY 1018 GAATATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTCAAGCAGGCTGCAA 1077
Db 161 GluTyThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuLyS 180
QY 1078 GATACTAAGCTATTTGAAAACATGACTTATGGGTGACCAATCACTTCAAGAAATTACTA 1137
Db 181 AspThrLySLeuLeuTyRThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeu 200
QY 1138 GCTCAAGCAAGCAATTTTAAACAAACCAAGCCAGCTATACGATTTTGAACGAGAC 1197
Db 201 AlaGlnAlaGlnSerIleLeuAsnLySAsnHisProGlyTyThrIleTyRGlunArgAsp 220

QY 1198 TCTCAATCGTCACTCATGACATGATTTTCCGTACGATTTTACCAATGATCAAGAG 1257
 DB 221 SerSerIleValThHIsapsbnaspIlephnArgThIleuProkethaspGIngu 240
 QY 1258 TTTACTACCGTGTAAATAATCGGAAACAGCTTTATGAGATCAATAAATAATCGTGTG 1317
 DB 241 PheThrTyArgValIleAsnArgInguInAlaTyArgIleAsnLysSerGIngu 260
 QY 1318 AATGAAGAATAAACAACCTGACCTGATCTCTGAGAAATATTACCTCTTAAAAAGG 1377
 DB 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGInguTyArgValIleuLysGly 280
 QY 1378 GAAAGCCGTATGATCCCTTTGATGACGAGTCACTGAAACCTGTCACATCAATGCTT 1437
 DB 281 GluLysProTyArgProPheAspArgSerHIsleuLysLeuPheThrIleuTyArgVal 300
 QY 1438 GATGTCGATACCAACGAATGCTTAAAGTAGACAGCTCTTAAACAGCTACGAACTTAAC 1497
 DB 301 AspValAspThrAsnGluLeuLeuLysSerGInguLeuThrIleAspGInguArgAsn 320
 QY 1498 TTAAGCTTCAAGATTTTATGATCTCTGATGATAGGCTTAACTTCAACAATCTTC 1557
 DB 321 LeuAspPheArgAspLeuTyArgProArgAspLysAlaIleLeuLeuTyArgAsnLeu 340
 QY 1558 GATGCTTTGGATTTATGAGACTATACCTTAACTGGAAGAGTAGAGATATACAGATGAC 1617
 DB 341 AspAlaPheGlyIleMetCaspTyThrLeuThrGlyLysValGluAspAsnHIsAspAsp 360
 QY 1618 ACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAAAGAGAGATGCTAGCTAT 1677
 DB 361 ThrAsnArgIleIleThrValTyMetGlyLysArgProGluGluLysAsnHIsAspTyArg 380
 QY 1678 CATTTAGCTTATGATTAAGATGTTATACCGAAGAAAGAGAGATTTACAGCTACCTG 1737
 DB 381 HisLeuAlaTyArgLysAspArgTyThrGluGluGluArgGluValTySerTyArgLeu 400
 QY 1738 CGTTATACAGGAGACCTATACCTGATTAACCTTAACGACAAA 1779
 DB 401 ArgTyThrGlyThrProIleProAspAsnProAspAspLys 414
 RESULT 11
 ADL92189
 ID ADL92189 standard; protein; 413 AA.
 AC ADL92189;
 XX 20-MAY-2004 (first entry)
 DT 20-MAY-2004 (first entry)
 DE Streptokinase protein sequence.
 XX harvesting; recombinant; host cell; N-terminal leader peptide;
 KW pre-peptide; lantibiotic; post-translational modification;
 KW pharmaceuticals; vaccine; immunogenic.
 OS Unidentified.
 XX W0200309862-A1.
 PN W0200309862-A1.
 PD 04-DEC-2003.
 XX 26-MAY-2003; 2003WO-NL000389.
 PF 24-MAY-2002; 2002EP-00077060.
 PR 07-FEB-2003; 2003US-00360101.
 XX (NANO-) APPLIED NANOSYSTEMS BV.
 PA Mol1 GN, Leenhouts CJ, Kuipers OP, Driessen AJM;
 PT WPI; 2004-042770/04.
 XX Harvesting a desired polypeptide produced by a recombinant host cell, for

PT producing pharmaceuticals, comprises selecting a recombinant nucleic acid
 PT comprising nucleic acid fragments encoding a leader peptide and the
 PT polypeptide.
 XX Claim 4; Page 82-83; 109pp; English.
 PS
 CC The invention relates to a novel method for harvesting a (poly)peptide
 CC produced by a recombinant host cell. The novel method involves selecting
 CC a cell comprising a first nucleic acid encoding a leader peptide and a
 CC second nucleic acid fragment encoding the desired (poly)peptide. The
 CC first and second fragments are within the same open reading frame of the
 CC first nucleic acid and the leader peptide is functionally equivalent to
 CC an N-terminal leader peptide found with the pre-peptide of a lantibiotic.
 CC The host cells and nucleic acids are useful for producing, harvesting and
 CC post-translational modification of polypeptides. The polypeptides may be
 CC used in the production of pharmaceuticals, e.g. as antigen for vaccine or
 CC immunogenic composition. This sequence represents a polypeptide relating
 CC to the novel method of the invention.
 CC
 SQ Sequence 413 AA;
 Alignment Scores:
 Pred. No.: 2,8e-192 Length: 413
 Score: 2153.00 Matches: 413
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 68.70% Indels: 0
 DB: 8 Gaps: 0
 US-09-940-235-11 (1-1782) x ADL92189 (1-413)
 QY 538 ATGCTGAGCCGTGAGGCGCTGATGACCGGCTCACTGTCACCAACAGCAATGCTTGT 597
 DB 1 IleAlaGlyProGluTyPheLeuLeuAspArgProSerAlaAsnSerGInguValVal 20
 QY 598 AGCGTTCGTGTCCTGTTGAGGGAAGCAATCAAGACATTAAGCTTAAATTTTGAATC 657
 DB 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPheGluIle 40
 QY 658 GATCTAACATCAGACCTGCTCATGAGAGAAAGAGAGAGAGCTTAAGTCCAAATCA 717
 DB 41 AspLeuThrSerArgProAlaHIsGlyGlyLysThrGluGluGluLeuSerProLysSer 60
 QY 718 AAACCATTTGCTACTGATAGTGGCGGATGTCATTAATTTGAGAAAGCTGACTTACTA 777
 DB 61 LysProPheAlaThrAspSerGlyAlaMetSerHIsLysLeuGluLysAlaAspLeu 80
 QY 778 AAGGCTATTCAAGAACAAATTGATCGCTAACGTCACAGTAAAGCGACTCTTGAAGTGC 837
 DB 81 LysAlaIleGlnGluGluLeuIleAlaAsnValHIsSerAsnAspAspTyArgGluVal 100
 QY 838 ATTGATTTTGAAGCGATGACCACTTACTGATGAAACGGCAAGGTCTACTTGTGCTGAC 897
 DB 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyValTyPheAlaAsp 120
 QY 898 AAAAGATGTTCCGTTAACTTGGCCGACCCCAACTGTCACAAAGATTTTGTCTAAGCGCAT 957
 DB 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGInguHIs 140
 QY 958 GTGCGGTTAAGACCATTAAGAAAAACCAATCAAAACCAAGGAAATGTTGAAGTGC 1017
 DB 141 ValArgValArgProTyArgGlyLysGlyLysProIleGlnAsnGlnAlaLysSerValAspVal 160
 QY 1018 GATATATGTCACAGTTTACTCCCTTAAACCTGATGACGATTTCAAGACGAGTCTCAAA 1077
 DB 161 GluTyThrValGlnThrThrProLeuAsnProAspAspAspPheArgProGlyLeuLys 180
 QY 1078 GATACTAAGCTATTGAAAAACATGATCGTGGAACCATCAATCATCTGAAGATTACTA 1137
 DB 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGInguLeuLeu 200
 QY 1138 GCTCAAGCAACAAAGCAATTTTAAACAAAAACCAACCGGCTATGATTAAGAACGAGAC 1197

Db 201 AlaGlnAlaGlnSerIleLeuAnlyAsnHisProGlyTyrThrIleTyrGluArgAsp 220
 QY 1198 TCCTCAATCGTCACTGATGCAATGACATTTTCGGTACGATTTTACCAATGATCAAG 1257
 Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
 QY 1258 TTTACTTACCGTGTAAAAATCGGAAACAAGCTTATAGATCAATAAAAAATCGTCTG 1317
 Db 241 PheThrTyrArgValIleAsnArgGluGlnAlaTyrArgIleAsnIleValSerGlyLeu 260
 QY 1318 AATGAAGAATAAACAACATGACCTGATCTCTGAGAAATATTCGCTTAAAAAGCG 1377
 Db 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluTyrTyrValIleuLysGly 280
 QY 1378 GAAACCCGATATATCCCTTGGATCGCATCTTGAACCGTTCACATCAATAGTT 1437
 Db 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheTrIleLysTyrVal 300
 QY 1438 GATGTCGATACCAACGAAATGCTAAAAAGTAGACAGCTTTAAACAGCTAGCGAACTTAC 1497
 Db 301 AspValAspThrAsnGluLeuLeuLysSerGluGlnLeuThrAlaSerGluArgAsn 320
 QY 1498 TTRGACTTCAGAGATTTATTCGATCTCTCGATAGGCTTAACTCTTACAAATCTC 1557
 Db 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
 QY 1558 GATGCTTTGGTATTTAGACTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTAC 1617
 Db 341 AspAlaPheIleLysLeuAspArgTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
 QY 1618 ACCAACCGTATCAATACCGTTTATATGAGGCAACGACCCGAGAGAGAGAAATGCTAGCTAT 1677
 Db 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProIleuGluAsnAlaSerTyr 380
 QY 1678 CATTTAGCTGATGATTAAGATGCTTATACCGAAAGAAAGAGAAAGTTTACAGTACTG 1737
 Db 381 HisLeuAlaTyrAspLysAspArgTyrThrGluGluGluArgGluValTyrSerTyrLeu 400
 QY 1738 CGTTATACGAGGACACCTTATACCTGATTAACCTTACGAC 1776
 Db 401 ArgTyrThrGlyThrProIleProAspAsnProAspAsp 413
 RESULT 12
 AAM21726
 ID AAM21726 standard; protein; 1194 AA.
 AC AAM21726;
 DT 17-OCT-2003 (revised)
 DT 01-OCT-1997 (first entry)
 XX Streptokinase/maltose binding protein fusion protein, rSK.
 DE Streptokinase/maltose binding protein fusion protein, rSK.
 KW Plasminogen-binding fragment; streptokinase; degradation; MBP; thrombolytic agent; blood clot; bolus; maltose-binding protein.
 XX Streptococcus dysgalactiae subsp. equisimilis.
 OS Streptococcus dysgalactiae subsp. equisimilis.
 XX
 XX Key Location/Qualifiers
 FH 1..381
 FT Protein /label= Maltose binding protein
 FT /note="acts as blocking group"
 FT Protein 382..1194
 FT /label= Streptokinase
 PN MO9641883-A1.
 XX 27-DEC-1996.
 PD 27-DEC-1996.
 XX 07-JUN-1996; 96WO-US009640.
 PF 07-JUN-1996; 96WO-US009640.
 XX 09-JUN-1995; 95US-00488940.
 PR 09-JUN-1995; 95US-00488940.
 XX

PA (HARD) HARVARD COLLEGE.
 XX
 XX Read GL;
 XX
 DR WPI; 1997-065469/06.
 XX
 PT Modified forms of streptokinase resistant to enzymatic cleavage - useful
 PT as thrombolytic agents in treating thrombosis and in medical equipment.
 XX
 XX Example 1, Page 11-12; 65pp; English.
 XX
 CC This sequence represents a fusion protein between maltose-binding protein
 CC and the plasminogen-binding fragment of streptokinase. This fusion
 CC protein was used in the design of modified streptokinase has an in vitro
 CC degradation rate at least 2 times slower than that of native
 CC streptokinase. Compounds containing modified streptokinases are
 CC specifically used as thrombolytic agents for dissolving blood clots in
 CC vivo in a mammal, preferably at a dose of 2000 U/kg, optionally as a
 CC bolus rather than by continuous infusion. (Updated on 17-OCT-2003 to
 CC standardise OS field)
 XX
 SO Sequence 1194 AA;
 Alignment Scores:
 Pred. No.: 7,46e-192 Length: 1194
 Score: 2150.50 Matches: 420
 Percent Similarity: 97.68% Conservative: 1
 Best Local Similarity: 97.45% Mismatches: 5
 Query Match: 68.62% Indels: 5
 DB: 2 Gaps: 3
 US-09-940-235-11 (1-1782) x AAM21726 (1-1194)
 QY 496 CAGACCAATCGAC-----GATCTGCGCCCTTACCGATGTTGATTTGCTGA 546
 Db 366 GlnThrAsnSerSerValProGlyArgGlySerIle---GluGlyArgIleAlaGly 384
 QY 547 CCTGAGTGGCTGCTTACACCGCTCCATCTGTCAACAACCAATGTTGTTAGCGTTGCT 606
 Db 385 ProGluTyrLeuLeuAspArgProSerValAsnAsnSerGlnLeuValIleSerValAla 404
 QY 607 GGTACTGTGAGGGAGCAATCAAGACATTAAGCTTAAATTTTGAATGATCTACA 666
 Db 405 GlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIleAspLeuThr 424
 QY 667 TCAAGACCTGCTCAATGAGGAAAGACAGAGCAAGCTTAACTCAAAATCAAAATCAATT 726
 Db 425 SerArgProAlaHisGlyGlyLysThrGluGlnIleuSerProLysSerLysProPhe 444
 QY 727 GCTAAGTATAGTGGCGGATGTCACATTAAGCTTGAAGAAGCTGACTTAAGGCTATT 786
 Db 445 AlaThrAspSerGlyAlaMetSerHisLeuGluLysAlaAspLeuLysAlaIle 464
 QY 787 CAAGAACAATGATCGGTAACGTCACAGTAACAGACAGTCTTGAAGGTCATGATTT 846
 Db 465 GlnGluGlnLeuIleAlaAsnValHisSerAsnAspArgTyrPheGluValIleAspPhe 484
 QY 847 GCAAGCATGCAACCATTAATGATCGAAGCGCAAGGCTTACTTGTGCAAGATAGGT 906
 Db 485 AlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAspLysGly 504
 QY 907 TCGGTAACTTGGCCGACCAACCTGTCCAAAGATTTTGTAGCGGACATGTGGCGTT 966
 Db 505 SerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHisValArgVal 524
 QY 967 AGACCAATTAAGAAAAACAATTAACAACCAAGCGAAATCTGTTGATGTGGAATTAATCT 1026
 Db 525 Arg---TyrLysGlyLysProIleGlnAsnGlnAlaLysSerValAspValGluTyrThr 543
 QY 1027 GTACAGTTTACTCCCTTAAACCTGATGACGATTTGACACAGGCTCAAGATATCAAG 1086
 Db 544 ValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLysAspThrLys 563

QY 1087 CTATTGAAACAGTACGATATCGGTGACACCATCATCTCAAGAACTTACTGACCAACA 1146
 Db 564 LeuLeuLysThrLeuAlaIleGlyAspThrIleThrsSerGlnGluLeuAlaGlnAla 583
 QY 1147 CAAGCATTTTAAACAAAACCAACCGGCTATACGATTTTAAAGTGAATCTTCAATC 1206
 Db 584 GlnSerIleLeuAsnLysAsnH1SPGGLYTYrThrIleTyrcIuArgAspSerSerIle 603
 QY 1207 GTCACTCATGACATGACATTTTCCGTACGATTTTACGAATTTGATGATGATGATTTTAC 1266
 Db 604 ValThrIleAspAsnAspIlePheArgThrIleLeuProMetCaspGlnGluPheThrTy 623
 QY 1267 CGTGTAAATAATCGGGACACAGCTTATAGATCATATAATAAAATCTGCTGATGAAAGA 1326
 Db 624 ArgValLysAsnArgGlnGlnAlaTyArgIleAsnLysSerGlyLeuAsnGlnGlu 643
 QY 1327 ATTAACACACCTGACCTGATCTCTGGAATAATTTACGCTCTTAAAGGGGAAAAGCCG 1386
 Db 644 IleAsnAsnThrAspLeuIleSerGlnLysTyTyValLeuLysGlyGlnLysPro 663
 QY 1387 TATGATCCCTTGTGATCGACGATCACTGAAACTGTTACACCATATACGTTGATGAT 1446
 Db 664 TyrAspProPheAspArgSerIleLeuLysLeuPheThrIleLysTyValAspValAsp 683
 QY 1447 ACCAAGCAATTCCTAAAGAGTGAAGAGCTCTTACAGCTAGCGAAGCTTACCTTAC 1506
 Db 684 ThrAsnGlnLeuLeuLysSerGlnGlnLeuLeuThrIleSerGlnArgAsnLeuAspPhe 703
 QY 1507 AGAGATTATACGATCTCTGTATGAAGGCTAAATCTACTTACCAACATCTCGATGCTTT 1566
 Db 704 ArgAspLeuTyAspProArgAspLysAlaLysLeuLysTyAsnAsnLeuAspAlaPhe 723
 QY 1567 GGTATTATGACTATATACCTTATCTGAAAGTGAAGAGAAATACGATGACCAACCT 1626
 Db 724 GlyIleMetAspTyThrLeuThrGlyLysValGluAspAsnH1AspAspThrAsnArg 743
 QY 1627 ATCAATACCGTTTATATGGAAGAGCAACCGAAGAGAGAAATGCTATCATTTAGCC 1686
 Db 744 IleIleThrValTyMetGlyLysArgProGlnGlnLysAsnH1AspSerTyThrIleAla 763
 QY 1687 TATGATTAAGATCGTTTATACGGAAGAAGAGAAAGTTTACAGCTACCTGCTTATACA 1746
 Db 764 TyrAspLysAspArgTyThrGlnGlnGluArgGlnValTySerTyLeuArgTyThr 783
 QY 1747 GGAGACCTATATACCTGATTAACCTTACGCAAA 1779
 Db 784 GlyThrProIleProAspAsnProAsnAspLys 794
 RESULT 13
 AAR12891
 ID AAR12891 standard; protein; 499 AA.
 XX
 AC AAR12891;
 XX
 AC 24-OCT-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 17-SEP-1991 (first entry)
 DE Streptokinase fused to yeast alpha factor secretion sequence.
 XX
 KM Fusion protein; blood clotting; coagulation; fibrinolysis;
 KM antithrombotic; thrombolysis; streptokinase.
 XX
 OS Streptococcus dysgalactiae subsp. equisimilis; ATCC 9542 or ATCC 100009.
 FT Key Location/Qualifiers
 FT Peptide 1..85
 FT /label= pre-pro alpha factor secretion sequence
 FT /note= "S. cerevisiae"
 FT Protein 86..499
 XX /label= mature streptokinase
 FN MO9109125-A.

XX 27-JUN-1991.
 PD
 XX 07-DEC-1989; 89GB-00027722.
 PF
 XX 07-DEC-1989; 89GB-00027722.
 PR 07-DEC-1990; 90MO-GB001911.
 XX
 PA (BRB1-) BRITISH BIO-TECHNOLOGY LTD.
 XX
 PI Dawson KM, Hunter MG, Czapliewsk LG;
 XX
 DR WPI; 1991-208151/28.
 DR N-PSDB; AAQ12158.
 XX
 PT Fusion protein cleavage by blood clotting enzyme - for prodn. of
 PT fractions having greater antithrombotic activity for therapy and
 PT prophylaxis.
 XX
 PS Disclosure; Page 86; 115pp; English.
 XX
 CC The streptokinase sequence was obt'd. from PCR amplified chromosomal DNA
 CC from S. equisimilis (Lancefield's Gp C) ATCC 10009 or ATCC 9642 (the
 CC primers used for the PCR were based on the published DNA sequence of S.
 CC equisimilis strain H46A (Malke, H., Roe, B., and Ferrer, J.D., Gene 34
 CC 357-362 (1985)). The gene was fused to DNA encoding the yeast alpha
 CC factor pre-pro-secretion sequence in an expression vector, pSMD1/152, for
 CC prodn. of recombinant strepto- kinase in S. cerevisiae strain BJ2168. See
 CC also AAR12887-R12889, AAR12892-R12894, AAR12885 and AAR12522. (Updated on
 CC 25-MAR-2003 to correct PA field.) (Updated on 24-OCT-2003 to standardise
 CC OS field)
 XX
 SQ Sequence 499 AA;
 Alignment Scores:
 Pred. No.: 2,62e-191 Length: 499
 Score: 2143.00 Matches: 413
 Percent Similarity: 96.29% Conservative: 2
 Best Local Similarity: 95.82% Mismatches: 16
 Query Match: 68.38% Indels: 0
 DB: 2 Gaps: 0
 US-09-940-235-11 (1-1782) x AAR12891 (1-499)
 QY 487 ACCTCTGTGACAGCCACATGACGCGATCTGCCCCCTTACCGATGTTGTTATGCTGCA 546
 Db 69 ThrIleAlaSerIleAlaIleAlaLysGlnGlnGlnValSerLeuAspLysArgIleAlaGly 88
 QY 547 CCTGAGTGGCTGTGTAACCGCTCATCTGTCAACAGCAAGCAATTTGTTAGCGTGTCT 606
 Db 89 ProGlnTrpLeuLeuAspArgProSerValAsnAspSerGlnLeuValValSerValAla 108
 QY 607 GGTACTGTGAGGGGAGCATCAAGACATTAGTCTTAAATTTTGAATCGATCTTACA 666
 Db 109 GlyThrValGlnGlnLysThrAsnGlnAspIleSerLeuLysPheGlnIleAspLeuThr 128
 QY 667 TCAGCACTGCTGTATGAGAGAAAGACAGCAAGCGCTTAAAGTCAAAATCAAAACCTTT 726
 Db 129 SerArgProAlaH1SerGlyLysTrpGlnGlnGlnLysSerProLysSerLysProPhe 148
 QY 727 GGTACTGATAGTGGGCGCATGTGCACTTAACTTGAAGAAAGTCACTTAAAGGCTATT 786
 Db 149 AlaThrAspSerGlyAlaMetProH1LysLysLeuGlnLysAlaAspLeuLysAlaIle 168
 QY 787 CAAGAACATTTGATTCGCTAAGCTGACAGTACAGACAGACTTACTTGAAGTCAATTTT 846
 Db 169 GlnGlnGlnLeuLeuIleAsnValHisSerAsnAspAspTyThrGlnValIleAspPhe 188
 QY 847 GCAAGCGATGCAACCATTTACTGATCAAGACGCAAGGCTTACTTGTCTGCAAGATGT 906
 Db 189 AlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyThrAlaAspLysAspGly 208
 QY 907 TCGGTAACTTGGCCAGCAACCACTGTCCAAAGATTTTGTAAAGCGGACATGTCCGCGTT 966


```

Db      ||| 209 SerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHisValArgVal 228
Qy      967 AGACCATATAAGAAAAACCAATATACAACCAAGCGAAATCTGTGATGGAATATACT 1026
Db      229 ArgProTylLeuSululysProIleGlnAsnGlnAlaIylsSerValAspValGluTyrThr 248
Qy      1027 GTRACGATTACTCCCTTAAACCTGATGACGATTTCAGACCAGGTCTCAAGATACTAG 1086
Db      249 ValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLysAspThrLys 268
Qy      1087 CTAATTGAAAACTAGTATCGGTGACACATGACATCTCAAGAACTACTAGCTCAGCA 1146
Db      269 LeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuAlaGlnAla 288
Qy      1147 CAAGCATTTTAAACAAAAACCAACCGGCTATACGATTATGAAACGTGACCTGCAATC 1206
Db      289 GlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluLysAspSerIle 308
Qy      1207 GTCACTGATGACATGACATTTTCCGTACGATTTTACCAATGATCAAGATTACTTAC 1266
Db      309 ValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGluPheThrTyr 328
Qy      1267 CGTGTTAAAAATCGGGAACAGCTTATAGATCAATAAAAATCGGTCTGATGAAAGAA 1326
Db      329 HisValLysAsnArgGluGlnAlaTyrGluIleAsnLysSerGlyLeuAsnGluGlu 348
Qy      1327 ATAAACAACAGTACCTGATCTCTGAGAAATATTAGCCTCTTAAAAAGGGAAACCG 1386
Db      349 IleAsnAsnThrAspLeuLysSerGlnLysTyrTyrValLeuLysLysGlyGluLysPro 368
Qy      1387 TATGATCCCTTGTGATGACAGTCACTTGAAACTGTCACCATCAATCTGTGATGTCAT 1446
Db      369 TyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrValAspValAsn 388
Qy      1447 ACCAAGCAATTCGTAAAAAGTAGAGAGCTCTTAAACAGTACCAACGTAACCTTAC 1506
Db      389 ThrAsnGluLeuLeuLysSerGlnGlnLeuThrAlaSerGluArgAsnLeuAspPhe 408
Qy      1507 AGAGATTTATAGATCTCTGTGATAAGGCTAAAGCTACTACCAACATCTCGATGCTTT 1566
Db      409 ArgAspLeuLysAspProArgAspLysAlaLysLeuLeuLysTrasnLeuAspAlaPhe 428
Qy      1567 GGTATTATGACATATCTTAACTGAGAAAGTAGAGATATACGATGACCAACCGT 1626
Db      429 GlyIleMetAspTyrThrIleThrGlyLysValGlnAspAsnHisAspAspThrAsnArg 448
Qy      1627 ATCATTAACCGTTATATGCGCAAGCGACCGGAAGAGAGATGCTAGTACATTTAGCC 1686
Db      449 IleIleThrValTylMetGlyLysArgProGluGlyGluAsnAlaSerTyrHisLeuAla 468
Qy      1687 TATGATTAAGATCGTTATATCCGAAGAAGAAAGAAAGTTTACAGCTACCTGCTTATACA 1746
Db      469 TyrAspLysAspArgTyrThrGlnGluGlnArgLysValTyrSerTyrLeuAsnArgTyrThr 488
Qy      1747 GGGACACCATATCTGATTAACCTTAAGCAAAA 1779
Db      489 GlyThrProIleProAspAsnProAsnAspLys 499

```

```

OS      Streptococcus dysgalactiae subsp. equisimilis.
XX
XX      PN      WO9641883-A1.
XX      PD      27-DEC-1996.
XX      PF      07-JUN-1996; 96WO-US0009640.
XX      PR      09-JUN-1995; 95US-00488940.
XX      PA      (HARD ) HARVARD COLLEGE.
XX      P1      Reed GL;
XX      DR      WPI; 1997-065469/06.
XX      PT      Modified forms of streptokinase resistant to enzymatic cleavage - useful
PT      as thrombolytic agents in treating thrombosis and in medical equipment.
XX
XX      PS      Example 1; Page 12-13; 65pp; English.
XX      CC      This sequence represents the wild type plasminogen-binding fragment of
CC      streptokinase. This fragment was used in the design of a modified
CC      streptokinase has an in vitro degradation rate at least 2 times slower
CC      than that of native streptokinase. Compounds containing modified
CC      streptokinases are specifically used as thrombolytic agents for
CC      dissolving blood clots in vivo in a mammal, preferably at a dose of 20000
CC      U/kg, opt. as a bolus rather than by continuous infusion. (updated on 17-
CC      OCT-2003 to standardise OS field)
XX
XX      SQ      Sequence 813 AA;
XX
XX      Alignment Scores:
XX      Pred. No.: 5.5e-191 Length: 813
XX      Score: 2140.50 Matches: 413
XX      Percent Similarity: 99.76% Conservative: 0
XX      Best Local Similarity: 99.76% Mismatches: 0
XX      Query Match: 68.30% Indels: 1
XX      DB: 2 Gaps: 1
XX
XX      US-09-940-235-11 (1-1782) x AAW21728 (1-813)
Qy      538 ATTGCTGACCGTGAAGTGGCTGCTAGACCGTCCATCTGTCACAAACAGCAATGGTGT 597
Db      1 IleAlaGlyProGluThrProLeuLeuAspArgProSerValAsnAsnSerGlnLeuValAl 20
Qy      598 AGCGTTGCTGTACTGTGAGGGGACGATCAAGACATTAAGCTTAAATTTTGAATC 657
Db      21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 40
Qy      658 GATTTAATCAGCAGCTGCTCATGAGAGAAAGACAGCAAGGCTTAAGTCAAAATCA 717
Db      41 AspLeuThrSerAspProAlaHisGlyGlyLysThrGlnGlnGlyLeuSerProLysSer 60
Qy      718 AAACCATTTGCTACTGATAGTGGCGGATGTCATATAACTTGAAGAAAGCTGACTTACTA 777
Db      61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 80
Qy      778 AAGGCTATTCAGAACCAATGATCGTACGTCACAGTAAAGCACTACTTGGAGTGC 837
Db      81 LysAlaAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100
Qy      838 ATTGATTTTCAAGCGAGTGAACCATTAAGTATGATGAAGCAAGGCAAGTCTACTTTGCTGAC 897
Db      101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
Qy      898 AAAGATGTTTCGGTGAACCTTGCCGACCAACCTGTCAGAAATTTTGGTAAAGCGACAT 957
Db      121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuSerGlyHis 140
Qy      958 GTGCGGCTTGAACCATTAAGAAACCAATACAAACCAAGGAATCGTTGATGAGTG 1017
Db      141 ValArgValArg--TyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 159

```

RESULT 14
 AAW21728
 ID AAW21728 standard; protein; 813 AA.
 XX
 AC AAW21728;
 XX
 DT 17-OCT-2003 (revised)
 DT 01-OCT-1997 (first entry)
 XX
 DE Wild type plasminogen-binding fragment of Streptokinase.
 KW Plasminogen-binding fragment; streptokinase; degradation; MBP;
 KW thrombolytic agent; blood clot; bolus; maltose-binding protein.

```

QY 1018 GAATATAGTACAGTTTACTCCCTTAAACCTGATGACGATTTCAGACGAGTCTCAAA 1077
DB 160 GUTYRTRHValGlnPheThrProLeuAenProAspAspAspPheArgProGlyLeuLys 179
QY 1078 GATCTAAGCTATTTGAAAACACTAGCTATCGGTGACACCATCACTCAAGAAATTACTA 1137
DB 180 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 199
QY 1138 GCTCAAGCAAGACGATTTTAAACAAAACCCAGGCTATTCGATTATGAAAGTGCAC 1197
DB 200 AlaGlnAlaGlnSerIleLeuLysnLysAsnHisProGlyThrThrIleThrGluAsp 219
QY 1198 TCCTCAATCGTCACTCATGACATGATCTTCCTGACGATTTTACCAATGATCAAGAG 1257
DB 220 SerSerIleValThrHisAspAsnAspIlePheAspThrIleLeuProMetAspGlnGlu 239
QY 1258 TTTACTTACCGTGTAAATAATCGGGAACAAGCTTATAGATCAATAATAATCTGCTTG 1317
DB 240 PheThrTrpArgValLysAsnArgGlnGluAlaTrpArgIleAsnLysLysSerGlyLeu 259
QY 1318 AATGAAGAATAATTAACAACACTGACCTGATCTCTGAGAAATATTACGCTTAAAAAGG 1377
DB 260 AsnGlnGlnIleAsnAsnThrAspLeuIleSerGlnLysTrpValLeuLysGly 279
QY 1378 GAAAAGCGTATGATCCCTTTGATGCGACGATCTTGAACCTGTCACCAATACGTT 1437
DB 280 GlnLysProTrpAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTrpVal 299
QY 1438 GATGTGATACCAACGAATTTGCTAAATAAGTGACAGCTCTTAAACAGTACGAACCTAAC 1497
DB 300 AspValAspThrAsnGlnLeuLeuLysSerGlnGlnLeuThrAlaSerGlnLysAsn 319
QY 1498 TTAGACTTCAGAGATTTATACATCTCTGATTAAGGCTTAACTCTTCAACAATCTC 1557
DB 320 LeuAspPheArgAspLeuLysAspProArgAspLysAlaLysLeuLeuTrpAsnAsnLeu 339
QY 1558 GATGCTTTTGGTATTTATGACATCTTAACTCTGAAAAGTGAGATATACCGATGAC 1617
DB 340 AspAlaPheGlyLysIleMetAspTrpThrLeuThrGlnGlyValGlnAspAsnHisAspAsp 359
QY 1618 ACCAACCGTATCATACCGTTTATATGCGCAAGCGACCGAAGAGAGAAATGCTAGCTAT 1677
DB 360 ThrAsnArgIleIleThrValTrpMetGlyLysArgProGlnGlnLysAsnAlaSerTrp 379
QY 1678 CATTTAGCTTATGATTAAGATCGTTATACCGAAGAAAGAGAGCTTTACAGCTTACTG 1737
DB 380 HisLeuAlaTrpAspLysAspArgTrpThrGlnGlnGlnLysValTrpSerTrpLeu 399
QY 1738 CGTTATACAGGAGACCTTATACCTGATTAACCTTAACGACGAAA 1779
DB 400 ArgTrpThrGlnGlnTrpProIleProAspAsnProAsnAspLys 413

```

RESULT 15
AAM86143
ID AAM86143 standard, protein; 414 AA.

AC AAM86143;
DT 17-OCT-2003 (revised)
DT 03-MAR-1999 (first entry)
DE Streptokinase (SK) protein sequence.
KW Non-immunogenic; epitope; T-cell; immunogenicity; immune system; SK;
KW immunogl bulin; therapeutic; streptokinase.
OS Streptococcus dysgalactiae subsp. equisimilis.
PN MO9852976-A1.
PD 26-NOV-1998.

```

PF 21-MAY-1998; 98MO-GB001473.
XX
PR 21-MAY-1997; 97GB-00010480.
PR 31-JUL-1997; 97GB-00016197.
PR 28-NOV-1997; 97GB-00025270.
PR 02-DEC-1997; 97US-0067235P.
PR 14-APR-1998; 98GB-00007751.
XX
PA (BIOV-) BIOVATION LTD.
XX
PI Carr FU;
XX
DR WPI; 1999-045301/04.
XX
PT Reducing immunogenicity of proteins - by modifying the amino acid
PT sequence of the protein to eliminate potential epitopes for T-cells of a
PT given species.
XX
PS Example 6; Fig 28; 77pp; English.
XX
CC The invention relates to a method for the production of non-immunogenic
CC proteins. The method comprises determining at least part of the amino
CC acid sequence of the protein; (b) identifying in the amino acid sequence
CC one or more potential epitopes for T-cells (T-cell epitopes) of the given
CC species; and (c) modifying the amino acid sequence to eliminate at least
CC one of the T-cell epitopes identified in step (b) thereby to eliminate or
CC reduce the immunogenicity of the protein when exposed to the immune
CC system of the given species. A method of analysing a pre-existing protein
CC to predict the basis for immunogenic responses is also provided. The
CC methods can be used particularly for reducing the immunogenicity of
CC immunoglobulins or therapeutic proteins, e.g. Streptokinase (SK). The
CC products can be used for diagnosis and therapy. The present sequence
CC represents the amino acid sequence of the SK protein. (updated on 17-OCT-
CC 2003 to standardise OS field)
XX
SQ Sequence 414 AA;

```

Alignment Scores:

Pred. No.:	2,6e-190	Length:	414
Score:	2132.00	Matches:	409
Percent Similarity:	99.03%	Conservative:	1
Best Local Similarity:	98.79%	Mismatches:	4
Query Match:	68.03%	Indels:	0
DB:	2	Gaps:	0

```

US-09-940-235-11 (1-1782) x AAM86143 (1-414)
QY 538 ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCAATGGTGT 597
DB 1 IleAlaGlyProGlnTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
QY 598 AGCGTGTGCTGTAATGTTGAGGGGACGAATCAAGACTTATAGTCTTAAATTTTGAATC 657
DB 21 SerValAlaGlyThrValGlnGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 40
QY 658 GATCTAATCATCAGACCTGCTCATGAGGAAAGACAGACGAAAGCTTAAAGTCCAAATCA 717
DB 41 AspLeuThrSerArgProAlaHisGlyGlyLeuThrGlnGlnGlnLysLeuSerProLysSer 60
QY 718 AAACGATTTGCTCTGATGATGAGGCGGATGTCACATTAACCTGGAAGAGCTTACTA 777
DB 61 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGlnLysAlaAspLeuLeu 80
QY 778 AAGGCTATTCAAGAACATTTGATCGCTTAACGTCACAGTAAACGACTTATGAGGTC 837
DB 81 LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspArgTrpPheGluVal 100
QY 838 ATTGATTTTCAAGACGATGCAACCATTTCTGATCGAAACGCAAGGCTTACTTTGCTGAC 897
DB 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTrpPheAlaAsp 120
QY 898 AAGATGCTTCGATTAACCTTGGCGAACCAACCTGTCACAGAAATTTTGTAGGAGCAT 957

```

```
Db      121 LysAspGlySerValThrIleuProThrGlnProValGlnGlnPheLeuSerGlyHis 140
QY      958 GTGCGGCTTAGACCATATAAAGAAAAACAATACAAACCAAGGAAATCTGTGATGTC 1017
Db      141 ValArgValArgProTyrIysGlnIysProIleGlnSerGlnAlaIysSerValAspVal 160
QY      1018 GAATATACCTGTACAGTTTACTCCCTTAACCTCGATGACGATTTGACACCAAGTCTCAA 1077
Db      161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuIys 180
QY      1078 GATCTAAGCTATTGAAACACATAGCTATGGTGACACACATCACTCAAGATTACTA 1137
Db      181 AspThrLysLeuLeuIysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeu 200
QY      1138 GCTCAGACCAAGACATTTTAAACAAACCAACCAAGGCTATACGATTTTATGAAGTGAC 1197
Db      201 AlaGlnAlaGlnSerIleLeuAsnIysThrHisProGlyTyrThrIleTyrGlnArgAsp 220
QY      1198 TCCTCAATGTCACATGACCAATGACATTTTCCGATCGATTTTACCAATGATCAAGAG 1257
Db      221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGln 240
QY      1258 TTACTTACCGGTGTAAATAATCGGAAACAGCTTATAGATCAATATAAAATCTGTCGTG 1317
Db      241 PheThrTyrHisValLysAsnArgGlnGlnAlaTyrGlnIleAsnIysSerGlyLeu 260
QY      1318 AATGAAGAATAAACAACACTGACCTGATCTCGAAGAAATTTACGTCCTTAAAAAAGGG 1377
Db      261 AsnGlnGlnIleAsnAsnThrAspLeuIleSerGlnIysTyrTyrValLeuIysLysGly 280
QY      1378 GAAAACCGTATGATCCCTTGTATGCGAGTCACTTGAACGTTCACCATCAATACGTT 1437
Db      281 GluIysProTyrAspProPheAspArgSerHisLeuIysLeuPheThrIleIysTyrVal 300
QY      1438 GATGTCGATACCAACGAATGTCTAAAGTAGACAGCTCTTAAACGCTACGGAACGTAA 1497
Db      301 AspValAsnThrAsnGlnLeuLeuIysSerGlnGlnLeuThrAlaSerGlnArgAsn 320
QY      1498 TTAGACTTCAGAGATTATATCGATCCTCGGATGAAGCTAACTACTCTACAAACAATCTC 1557
Db      321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
QY      1558 GATGCTTTGGTATTATGACACTATACCTTAACTGAAAAGTAGAGATATACGATGAC 1617
Db      341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGlnAspAsnHisAspAsp 360
QY      1618 ACCAACCGTATCATAAACGTTTATATGCGGCAAGCAGCCGAGAGAGAGAAATGCTAGCTAT 1677
Db      361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGlnGlnIleAsnAlaSerTyr 380
QY      1678 CATTTAGCCTTAGATATAAGATGCTTATACCGAAGAGAGACGAGAAGTTTACAGCTACCTG 1737
Db      381 HisLeuAlaTyrAspLysAspArgTyrThrGlnGlnIleArgGlnValTyrSerTyrLeu 400
QY      1738 CGTTATACAGGAGACACTATACCTGATTAACCTTAACGACAA 1779
Db      401 ArgTyrThrGlyThrProIleProAspAsnProAsnAspLys 414
```

Search completed: January 28, 2006, 02:08:39
Job time : 145.96 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 28, 2006, 01:49:55 ; Search time 25.4212 Seconds
(withon alignments)
11590.966 Million cell updates/sec

Title: US-09-940-235-11

Perfect score: 3134

Sequence: 1 tgcgttcacgtcgtcgtcgcg.....ataccctaagcacaataa 1782

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 1144120

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

MODEL=frame+ n2p.model -DEV=rlp
-O=/cgn2_1/USFTO_spool_p/US09940235/runat_27012006_144219_27602/app_query.fasta_1.7708
-DB=Issued Patents_AA -QFMT=fastan -SUPFIX=rai -MINMATCH=0.1 -LOOPCL=0
-LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09940235 @CGN_1_1_159 @runat_27012006_144219_27602 -NCPU=6 -ICPU=3
-NO MMAP -LARGEOUTERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents_AA:*
1: /cgn2_6/prodata/1/iaa/5.COMB.pcp:*
2: /cgn2_6/prodata/1/iaa/6.COMB.pcp:*
3: /cgn2_6/prodata/1/iaa/H.COMB.pcp:*
4: /cgn2_6/prodata/1/iaa/PCTUS.COMB.pcp:*
5: /cgn2_6/prodata/1/iaa/RA.COMB.pcp:*
6: /cgn2_6/prodata/1/iaa/backfile1.pcp:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2168	69.2	795	2	US-09-211-542A-2
2	2165	69.1	483	1	US-07-854-596B-43
3	2158	68.9	414	2	US-09-211-542A-6
4	2158	68.9	414	6	5240845-1
5	2158	68.9	440	1	US-08-488-940-1
6	2153	68.7	413	2	US-10-360-101-264
7	2150.5	68.6	1194	1	US-08-488-940-1
8	2143	68.4	499	1	US-07-854-596B-28
9	2140.5	68.3	813	1	US-08-488-940-3
10	2132	68.0	415	1	US-07-854-596B-26
11	2132	68.0	435	1	US-07-854-596B-19
12	2132	68.0	440	1	US-07-854-596B-15

13	2132	68.0	483	1	US-07-854-596B-47	Sequence 47, Appl
14	2132	68.0	859	1	US-07-854-596B-35	Sequence 35, Appl
15	2130.5	67.7	1194	1	US-08-488-940-17	Sequence 17, Appl
16	2114.5	67.5	1194	1	US-08-488-940-18	Sequence 18, Appl
17	2090.5	66.7	413	1	US-08-759-599-12	Sequence 12, Appl
18	2090.5	66.7	413	2	US-09-294-457-12	Sequence 12, Appl
19	2090.5	66.7	413	2	US-09-919-703-12	Sequence 12, Appl
20	2071	66.1	800	1	US-08-488-940-4	Sequence 4, Appl
21	2071	66.1	1181	1	US-08-488-940-2	Sequence 2, Appl
22	2061	65.8	401	2	US-09-374-038-1	Sequence 1, Appl
23	2061	65.8	401	2	US-09-658-179-1	Sequence 1, Appl
24	2061	65.8	413	2	US-09-374-038-2	Sequence 2, Appl
25	2061	65.8	413	2	US-09-658-179-2	Sequence 2, Appl
26	1902	60.7	384	2	US-09-374-038-4	Sequence 4, Appl
27	1902	60.7	384	2	US-09-658-179-4	Sequence 4, Appl
28	1898	60.6	372	2	US-09-374-038-3	Sequence 3, Appl
29	1898	60.6	372	2	US-09-658-179-3	Sequence 3, Appl
30	1876	59.9	747	1	US-07-854-596B-40	Sequence 40, Appl
31	1875	59.8	369	1	US-07-854-596B-31	Sequence 31, Appl
32	1872.5	59.7	736	2	US-09-211-542A-4	Sequence 4, Appl
33	1860	59.3	356	2	US-09-211-542A-12	Sequence 12, Appl
34	785	25.0	150	2	US-09-211-542A-14	Sequence 14, Appl
35	504.5	16.1	2324	1	US-08-283-857-1	Sequence 1, Appl
36	504.5	16.1	2324	4	PCT-US95-09819-1	Sequence 1, Appl
37	504.5	16.1	2446	1	US-08-551-356-2	Sequence 2, Appl
38	504.5	16.1	2446	4	PCT-US93-12687-2	Sequence 2, Appl
39	503	16.0	2231	1	US-08-153-799-16	Sequence 16, Appl
40	503	16.0	2335	2	US-10-360-101-235	Sequence 235, App
41	503	16.0	2386	1	US-09-016-366A-12	Sequence 12, Appl
42	503	15.9	2327	6	US-09-961-403-1	Sequence 1, Appl
43	499.5	9.0	283	2	5455158-1	Parent No. 5455158
44	283	8.8	286	2	US-09-078-091-4	Sequence 4, Appl
45	275	8.8	286	2	US-09-078-091-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-09-211-542A-2
Sequence 2, Application US/09211542A
Patent No. 6210667
GENERAL INFORMATION:
APPLICANT: Reed, Guy L.
TITLE OF INVENTION: BACTERIAL FIBRIN-DEPENDENT PLASMINOGEN ACTIVATOR
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROMBERG & SUNSTEIN, LLP
STREET: 125 Summer Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/211,542A
FILING DATE: 15-December-1998
CLASSIFICATION: 1653
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/069,497
FILING DATE: 15-December-1997
ATTORNEY/AGENT INFORMATION:
NAME: Attorney, Strimpel, Harriet M.
REGISTRATION NUMBER: 37,008
REFERENCE/DOCKET NUMBER: 1874/111
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)443-9292
TELEFAX: (617)443-0004
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:

LENGTH: 795 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-211-542A-2

Alignment Scores:

Alignment No.:	Score:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
5, 2e-211	2168.00	795	421	1	5	2	2
Percent Similarity:	97.918						
Best Local Similarity:	97.688						
Query Match:	69.184						

US-09-940-235-11 (1-1782) x US-09-211-542A-2 (1-795)

```

QY 496 CAGACCAATCGAGC-----GGATCTGGCCCTTCACCGATGTCGTATGCTGGA 546
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 366 GlnThrAsnSerSerValProGlyArgGlySerIle--GlnGlyArgIleAlaGly 384

QY 547 CCTGAGTGGCTGCTAAGCCGTCATCTGTCAACACAGCCAAATTGGTTGTTACCGTTCCT 606
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 385 ProGlnTrpLeuAsnAspArgProSerValAsnAsnSerGlnLeuValValSerValAla 404

QY 607 GGATACGTTGAGGGGACGATCAAGACATTAGCTTAAATTTTGAATTCGATCTACACA 666
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 405 GlyTrpValGlnGlyThrAsnGlnAspIleSerLeuLysPhePheGlnIleAspLeuThr 424

QY 667 TCACGACCTGCTCATGAGAGAAAGACAGACGAGGCTTAAAGTCCAAATCAAAACGATT 726
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 425 SerArgProAlaHisGlyGlySerThrGlnGlnGlyLeuSerProLysSerLysProPhe 444

QY 727 GCTACTGATGATGGCCGCGCATGTCACATAAATTGAGAAAGCTGACTTAAAGCTATT 786
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 445 AlaTrpAspSerGlyAlaMetSerHisLysLeuGlnLysAlaAspLeuLysAlaIle 464

QY 787 CAAGAACAATTGATGCTAAGCTCCACGATTAACGACGACTTGTAGGTCATTGATTTT 846
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 465 GlnGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTrpPheGlnValIleAspPhe 484

QY 847 GCAGCGATGACACCAATTCTGATCGAAACGCGAAGGTCCTACTTGTGACAAAGATGAT 906
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 485 AlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAspLysAspGly 504

QY 907 TCGGTAACCTTGCAGCCGACCTGTCCAAAGAAATTTTGTCTAAGCGACATGTGCCGCT 966
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 505 SerValThrLeuProThrGlnProValGlnGlnPheLeuLeuSerGlyHisValArgVal 524

QY 967 AGACCATTAAGAAAAACCAATPACAAACCAAGCAATCTGTGATGCGAATATATCT 1026
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 525 ArgProTrpLysGlnLysProIleGlnAsnGlnAlaLysSerValAspValGlnTyrThr 544

QY 1027 GTACAGTTTACTCCCTTAAACCTGATGACATTTTCAAGACAGGCTCTCAAGATPACTAAG 1086
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 545 ValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuLysAspThrLys 564

QY 1087 CTATTGAAAAACACTAGCTATCGGTGACACCAATCAATCTCAAGAAATTTACTGCTACAGCA 1146
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 565 LeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeuLeuAlaGlnAla 584

QY 1147 CAAGACATTTTAAACAAAAACCAAGGCTATACATTTTANGAAGTACTCTTCAATC 1206
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 585 GlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGlnLysAspSerSerIle 604

QY 1207 GTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGATCAAGAGTTTACTTAC 1266
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 605 ValThrHisAspAsnAspIlePheAspGlnThrIleLeuProMetAspGlnLysPheThrTyr 624

QY 1267 CGTGTTAAAAATCGGGAAACAGCTTATAGATCAATTAATAAATCTGGTCTGAATGAAGA 1326
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 625 ArgValLysAsnArgGlnGlnAlaTyrArgIleAsnLysLysSerGlyLeuAsnGlnGln 644

QY 1327 ATAAACAACACATGACCTGATCTCTAGGAATAATTACGCTTAAAAAAGGGAAAAAGCCG 1386

```

```

Db 645 IleAsnAsnThrAspLeuIleSerGlnLysTyrTyrValLeuLysLysGlnLysPhe 664

QY 1387 TATGATCCCTTTGATCGCAGTCACTGAAACCTGTTCACCATCAATACGTTGATGTCAT 1446
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 665 TyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrValAspValAsp 684

QY 1447 ACCAAGCAATTTGCTAAATAAAGAGACGCTCTTAAACACTGACGCGAAGCTTACTGACTTC 1506
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 685 ThrAsnGlnLeuLeuLysSerGlnGlnLeuLeuThrAlaSerGlnLysAsnLeuAspPhe 704

QY 1507 AGAGATTTATACGATCCCTCGTATAGGCTTAAACTTACTTACAACTCTGATGCTTTT 1566
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 705 ArgAspLeuTrpAspProAlaGlyAspLysAlaLysLeuLeuLysTrpAsnAsnLeuAspAlaPhe 724

QY 1567 GGATTTATGACATPACCTTAACTGAGAAAGTAAAGGATATACGATGACACCAACCGT 1626
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 725 GlyIYIleMetAspTrpThrLeuThrGlnLysValGlnAspAsnHisAspAspThrAsnArg 744

QY 1627 ATGATAACGCTTTATATAGGCAAGCCGACCGAAGAGAGAAATGCTACTATCATTTAGCC 1686
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 745 IleIleThrValTyrMetGlyLysArgProGlnGlyGlnLysAlaSerTyrHisLeuAla 764

QY 1687 TATGATTAAGATCGTTATACCGAAGAAAGACGAAATTTACAGTCACTGCGTTATACA 1746
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 765 TyrAspLysAspArgTyrThrGlnGlnGlnLysGlnValTyrSerTyrLeuArgTyrThr 784

QY 1747 GGGACACTTATACCTGATTAACCTTAACGACAAA 1779
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 785 GlyThrProIleProAspAsnProAsnAspLys 795

RESULT 2
US-07-854-596B-43
; Sequence 43, Application US/07854596B
; Patent No. 5434073
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith M
; APPLICANT: Hunter, Michael G
; APPLICANT: Czaplowski, Lloyd G
; TITLE OF INVENTION: Proteins and nucleic acids
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Dr. John J. McDonnell
; STREET: Ten South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/854,596B
; FILING DATE: 03-JUN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,337
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SBO ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 483 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-854-596B-43

```

Alignment Scores:

Pred. No.:	8.15e-211	length:	483
Score:	2155.00	Matches:	434
Percent Similarity:	85.22%	Conservative:	10
Best Local Similarity:	83.30%	Mismatches:	27
Query Match:	69.08%	Indels:	50
DB:	1	Gaps:	9

US-09-940-235-11 (1-1782) x US-07-854-596B-43 (1-483)

QY	244	ACAGATTGTACCCCAATGAGCAAGAGTTTGTATCAATGCTGCTGGGACCTTCATGAGGT	303
Db	4	ThrsapCy ⁺ ethGlusercI.ygIn-----AsnLeucYs-----LeucYs---	16
QY	304	CGAGAAACGTGGAGAGGACAGCCGAGCCATCTTGCACTTCTAGAAATAGATGCAC	363
Db	17	-----GlnGlySerAsn-----ValCysGlyGlnGlyAsnYsCys---	28
QY	364	GATCAGACACAAAGGACATCCATATGAAATTGGAGACACCTGGAGCAAGAAAGTAATCGA	423
Db	29	-----::: -----IleuGlySerAspGlyGlnYsAsn-----	37
QY	424	CGAAACCTGCTCAGTGCATCTGCACAGGCAACGGCGGAGAGAGTGAAGTGTGAGAG	483
Db	38	-----GlnCysVal-----ThrGlyGlnGly-----	44
QY	484	CACACCTCTGTCAGACCAATCGAGCGGATCTGGCCCTTCACCGATTT-----	534
Db	45	-----ThrProYsProGlnSerHisAsnAspGlyAspPheGlnGlnIleProGlnGln	62
QY	535	-----CGATTGGTGAACCTGATGGCTGCTTAACCGTCCATCTGTC	576
Db	63	TyrleuGlnIleGlnGlyArgIlealaGlyProGlnIleuLeuAspArgProSerVal	82
QY	577	AACAAACGCCAATTGGTTGATACCGTGTGGTACTGTGAGGGAGCAATCAAGCAATT	636
Db	83	AsnAsnSerGlnLeuValIalSerValalaGlyThrValGlnGlyThrAsnGlnAspIle	102
QY	637	AGCTCTAAATTTTGGAAATCGATCTTAACATCAACGACTGCTCATGAGAGAAAGACAG	696
Db	103	SerleuYsPhePheGlnIleAspLeuThrSerArgProAlaHisGlyGlySerThrGln	122
QY	697	CAAGGCTTAAGTCCAAATCAAAACCACTTGCTACTGATAGTGGCCGATGTCACTTAA	756
Db	123	GlnGlyLeuSerProYsSerIysProPheAlaThrAspSerCylValaMetProHisYs	142
QY	757	CTTGAGAAAGCTGACTTACTAAAGGCTATTCAGAAACAATTGATGGCTAACGTCACAGT	816
Db	143	LeuGlnYsAlaAspLeuLeuYsAlaIleGlnGlnGlnLeuIleAlaAsnValHisSer	162
QY	817	AACGACGACTACTTGAGTCAATTGATTTGGACGGATGCAACCACTTACTGATCGAAAC	876
Db	163	AsnAspAspTyrPheGlnValIleAspPheAlaSerAspAlaThrIleThrAspArgAn	182
QY	877	GGCAAGGCTACTTTGCTGCAAAAGATGTTGGTAACTTGGCCGACCACTGTTCGAA	936
Db	183	GlyYsValTyrPheAlaAspYsAspGlySerValThrLeuProThrGlnProValGln	202
QY	937	GAATTTTGTCAAGGCGCATGTGGCGGCTTAACCATTAAGAAAGAAACCAATPACAAC	996
Db	203	GlnPheLeuLeuSerGlyHisValaArgValaArgProTyrLysGlnYsProIleGlnAn	222
QY	997	CAAGCGAAATCTGTGATGTGAATATATCTGACGTTACTCCCTTAAACCTGATGAC	105
Db	223	GlnAlaYsSerValAspValGlnTyrThrValGlnPheThrProLeuAsnProAspAsp	242
QY	1057	GATTTCAGACCAAGGCTCAAAAGATATCAAGCTATTGAAAAACTAGCTATCGGTGACCC	111
Db	243	AspPheArgProGlyLeuYsAspThrYsLeuLeuYsThrLeuAlaIleGlyAspThr	262
QY	1117	ATGCATCTCAAGAATTACTAGCTCAAGACAAAGATTTTAAACAAGAACCAACCAAGCC	117
Db	263	IleThrSerGlnGlnLeuLeuAlaGlnAlaGlnSerIleLeuAsnYsThrHisProGly	282

QY	1177	TATACATTTATGAACGTGACTCCTCATGCTGACTCATGACATTCGATGCGTACG	1236
Db	283	TyrThrIleTyrGluAgaApsSerIleValThrHisAspAenApeIlePheArgThr	302
QY	1237	ATTTTACCAATGATGATCAAGAGTTTACTTACCGGTATAAAATCGGAAACAAGCTTAGG	1296
Db	303	IleLeuProIleAepGInGluPheThrTyrHisValIysAsnAArgGluGlnAlaTyrGlu	322
QY	1297	ATCAATATAAAATCTGGTCTGGAAGAAGAAATAACAACCTGACCTGATCTTCGAGAA	1356
Db	323	IleAsnIlySylsSerGlyLeuAsnGluGlnIleAsnAsnThrApeIleIleSerGluLys	342
QY	1357	TATACGCTCCCTTAAAAAAGGGGAAACCCGATATATCCCTTGAATCCGACTTGA	1416
Db	343	TyrTyrValIleuIySylSglYglIulYerProTyrAspProPheAspApgSerHisIleuLys	362
QY	1417	CTGTTCACCAATCAATACGTTGATGTCGATACCAAGAAATTTCTAAAAAGTGACAGCTC	1476
Db	363	LeuPheThrIleuYerTyrValAAspValAsnThrAsnGluLeuIySerGluGlnIleu	382
QY	1477	TTTACACTGACCGAAGCTTACTTAGACTTCAGAGATTATACGATCCTCTCGATTAAGCT	1536
Db	383	LeuThrIleAspGluAgaAenLeuAspPheArgAspIeuYrAspProAgaAspLysAla	402
QY	1537	AAACTACTTACAAACAATCTCGATGCTTTGGTATTTAGACTATACCTTAACCTGAGAAA	1596
Db	403	LysIleuIeuYrIleAsnAenLeuAspAlaPheGlyIleIleAspTyrThrIleuThrGlyLys	422
QY	1597	GTAAGAGATATACGATGACACCAACCGTATCATACCGTTATATATGAGGCAAGCACCC	1656
Db	423	ValGluAspAsnHisAspAAspThrAsnAArgIleIleThrValTyrMetGlyLysArgPro	442
QY	1657	GAGGAGAAGAATGTGCTATCATTTAGCCTATGACATAAGATCGTTATACCGAAGAA	1716
Db	443	GluGlyGluAsnAspIleAspTyrHisIleuAlaTyrAspLysAspArgTyrThrGluGluGln	462
QY	1717	CGAAGATTATACAGCTTACTGCGTTATACAGAGGACACTTATACCTGATTAACCTTAACGAC	1776
Db	463	ArgGluValIyIleSerTyrIleuAArgTyrThrGlyThrProIleProAspAenProAsnAsp	482
QY	1777	AAA 1779	
Db	483	lys 483	
RESULT 3			
US-09-211-542A-6			
Sequence 6, Application US/09211542A			
Patent No. 6210667			
GENERAL INFORMATION:			
APPLICANT: Reed, Guy L.			
TITLE OF INVENTION: BACTERIAL FIBRIN-DEPENDENT PLASMINOGEN ACTIVATOR			
NUMBER OF SEQUENCES: 14			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: BROMBERG & SUNSTEIN, LLP			
STREET: 125 Summer Street			
CITY: Boston			
STATE: Massachusetts			
COUNTRY: USA			
ZIP: 02110			
COMPUTER READABLE FORM:			
MEDIUM TYPE: Floppy disk			
COMPUTER: IBM PC compatible			
OPERATING SYSTEM: PC-DOS/MS-DOS			
SOFTWARE: PatentIn Release #1.0, Version #1.25			
CURRENT APPLICATION DATA:			
APPLICATION NUMBER: US/09/211,542A			
FILING DATE: 15-December-1998			
CLASSIFICATION: 1653			
PRIOR APPLICATION DATA:			
APPLICATION NUMBER: 60/069,497			
FILING DATE: 15-December-1997			
ATTORNEY/AGENT INFORMATION:			

NAME: Attorney, Strimpel, Harriet M.
REGISTRATION NUMBER: 37,008
REFERENCE/DOCKET NUMBER: 1874/111
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)443-9292
TELEFAX: (617)443-0004
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 414 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-211-542A-6

Alignment Scores:
Pred. No.: 3,87e-210 Length: 414
Score: 2158.00 Matches: 414
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 68.86% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-11 (1-1782) x US-09-211-542A-6 (1-414)

QY 538 ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAAGCAAGCAATTGGTT 597
DB 1 IleaaglyProglutThrleuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
QY 598 AGCGTGTGCTGACTGTGTGAGGGAAGAAATCAAGACATTAGCTTAAATTTTGAATC 657
DB 21 SerValaIaaglyThrValGlnGlnThrAsnGlnAspIleSerLeuysPhepneGluile 40
QY 658 GATCTAACATCAAGACCTGCTCTAGAGGAAGAAGACAGAGCAAGCTTAAAGTCCAAATCA 717
DB 41 AspleuThrSerArgProAlaHisGlyGlyThrGlnGlnGlyLeuSerProlysSer 60
QY 718 AAACATTGTGCTACTGATGTGGCGGAGATGTCAATTAATCTGAGAAAGCTGACTACTA 777
DB 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluValAspLeuLeu 80
QY 778 AAGGCTATTCAAGAACATTTGCTAAGCTCAACAGTACAGACACTTCTTGAAGTC 837
DB 81 LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspArgPyrPheGluVal 100
QY 838 ATTGATTTTGGCAAGCGATCAACCATTTACTGATCAAGCGCAAGGCTTACTTGTGAC 897
DB 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyValValTyrPheAlaAsp 120
QY 898 AAAGATGGTGGCTAACTTGGCGAAGCCGATGTCAGAAATTTTGGTAAAGCGACAT 957
DB 121 LysAspGlySerValThrLeuProThrGlnProValGlnGlnPheLeuLeuSerGlyHis 140
QY 958 GTGCGCTTGAACCATTAAGAAAAACCAATACAAACCAAGCAAGAAATCTGTGATGTG 1017
DB 141 ValArgValArgProTyrIleGlySerProIleGlnAsnGlnAlaLysSerValAspVal 160
QY 1018 GAATTAAGTGTACAGTTTACCTCTTAAACCTGTATGACGATTTCAGACCAGGCTCAAA 1077
DB 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
QY 1078 GATTAAGCTTAAGTAAAAACATAGCTATCGGTGACACCATCACTCAAGAAATTAACTA 1137
DB 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeuLeu 200
QY 1138 GCTCAAGACAAGACATTTTAAACAAAACCAAGCCAGGCTATACGATTTATGAACGTAC 1197
DB 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGlnLysAsp 220
QY 1198 TCTTCAATGCTACTCATGACAAATGACATTTTCCGTAGCATTTTACCAATGATCAAGAG 1257
DB 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
QY 1258 TTTACTTACCGGTAAATAATGGGAACAAGCTTATAGATCAATAAAAATCTGGTCTG 1317

DB 241 PheThrTyrArgValLysAsnArgGlnGlnAlaTyrArgIleAsnLysSerGlyLeu 260
QY 1318 AATGAGAATTAACAACACTGACCTGATCTCTGAGAAATTTAGCTTTAAAAAGGG 1377
DB 261 AsnGlnGlnIleAsnAsnThrAspLeuIleSerGlnLysTyrValLeuLysGly 280
QY 1378 GAAAAGCCGATATGCTTGTGATCGGACGACCTTGAACGTTACCATCAATTAAGTT 1437
DB 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
QY 1438 GATGCTGATACCAACGATTTGCTTAAAGTAGAGCAAGCTTTAAGCTAGCGAAGTAAAC 1497
DB 301 AspValAspThrAsnGlnGlnLeuLeuLysSerGlnGlnLeuThrAlaSerGlnLysAsn 320
QY 1498 TTAGACTTCAGAGATTATTAACATCTCTGTATAGGCTTAACTTCAACAATCTC 1557
DB 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
QY 1558 GATGCTTTGGTATTTATGACCTATACCTTAACTGGAAGATGAGATTAACAGATGAC 1617
DB 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyValGlnAspAsnHisAspAsp 360
QY 1618 ACCAAGCGTATCAATACCGTTTATATGCGCAAGCAAGCCGAAGAGAGAAATGCTAGCTAT 1677
DB 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGlnGlnLysAlaSerTyr 380
QY 1678 CATTAGCTTATGATTAAGATCGTTATACCGAAGAAAGACAGAGATTACAGTACTG 1737
DB 381 HisLeuAlaTyrAspLysAspArgTyrThrGlnGlnGlnValTyrSerTyrLeu 400
QY 1738 CGTTATACAGGACACCTTATACCTGATTAACCTTAACGACAA 1779
DB 401 ArgTyrThrGlyThrProIleProAspAsnProAsnAspLys 414

RESULT 4
5240845-1
Patent No. 5240845
APPLICANT: FUJII, SETSURO; TAKADA, KAORUKO.; KATANO, TAMIKI;
MURATA, BIJUNO; KOICHI, ONO; KENJI, SAKATA, YASUYO; UENOYAMA,
TSUTOKU
TITLE OF INVENTION: MUTATED STREPTOKINASE PROTEINS
NUMBER OF SEQUENCES: 65
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/549,049
FILING DATE: 06-JUL-1990
SEQ ID NO: 1:
LENGTH: 414
5240845-1

Alignment Scores:
Pred. No.: 3,87e-210 Length: 414
Score: 2158.00 Matches: 414
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 68.86% Indels: 0
DB: 6 Gaps: 0

US-09-940-235-11 (1-1782) x 5240845-1 (1-414)

QY 538 ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAAGCAAGCAATTGGTT 597
DB 1 IleaaglyProglutThrleuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
QY 598 AGCGTGTGCTGACTGTGTGAGGGAAGAAATCAAGACATTAGCTTAAATTTTGAATC 657
DB 21 SerValaIaaglyThrValGlnGlnThrAsnGlnAspIleSerLeuysPhepneGluile 40
QY 658 GATCTAACATCAAGACCTGCTCTAGAGGAAGAAGACAGAGCAAGCTTAAAGTCCAAATCA 717
DB 41 AspleuThrSerArgProAlaHisGlyGlyThrGlnGlnGlyLeuSerProlysSer 60
QY 718 AAACATTGTGCTACTGATGTGGCGGAGATGTCAATTAATCTGAGAAAGCTGACTACTA 777

```

Db      |||
61      LysProPheAlaThrAspSerGlyAlaMetSerHisIleuLeuGluValAspLeuLeu 80
Qy      778 AAGGCTATTCAAGAACATTTGATCGCTAACGTCACAGTAACGACGATCTTTGAGTGC 837
Db      81 LysAlaIleGlnGluIleuIleAlaAsnValHisSerAsnAspIlePheGluVal 100
Qy      838 ATTGATTTTGGCAAGCGATGCAACCATTTACTGATCGAAACGGCAGGCTCTATTGCTGAC 897
Db      101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyValIlePheAlaAsp 120
Qy      898 AAAGATGCTTGGGTAACTTCCGACCCCAACTGTCACAGAAATTTTGTCTAACGGACAT 957
Db      121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
Qy      958 GTGCGCGTTAGACCATATTAAGAAAAACAATGACAAACCAACGAGAAATCTGTGATGTC 1017
Db      141 ValArgValArgProIleGlyGluValProIleGlnAsnGlnAlaLysSerValAspVal 160
Qy      1018 GAATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTGACAGCGAGTCTCAAA 1077
Db      161 GluIleThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
Qy      1078 GATACTTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCTCAAGATTACTA 1137
Db      181 AspThrLysLeuLeuThrLeuAlaIleGlyAspThrIleThrSerGlnIleuLeu 200
Qy      1138 GCTCAAGCACAAGACATTTTAAACAAAACACCCGCGTATACGATTATGACGTCAC 1197
Db      201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyIleThrIleIleGluAsp 220
Qy      1198 TCCCTCAATCGTCACTCATGACATGATTTCCGTACGATTTTACCAATGATCAAG 1257
Db      221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
Qy      1258 TTTACTTACCGTGTAAATAATCGGGAACAAGCTTATAGATCAATAATAATCGTGTG 1317
Db      241 PheThrIleArgValLysAsnArgGlnGlnAlaIleArgIleAsnLysLysSerGlyLeu 260
Qy      1318 AATGAAGAAATTAACAACACTGACCTGATCTCTGAGAAATATTATTCCTTAAAAAGG 1377
Db      261 AsnGlnGluIleAsnAsnThrAspLeuIleSerGluLysIleTyrlValLeuLysGly 280
Qy      1378 GAABAAGCGGTATGATCCCTTATATGCAATCACTGAACTGTAACATCAATACGT 1437
Db      281 GluLysProIleArgProPheAspArgSerHisLeuLysLeuPheThrIleLysIleVal 300
Qy      1438 GATGTCGATACCAAGAAATTTGCTAAATAAGTAGACAGCTCTTAAACAGCTAGCGAACTAAC 1497
Db      301 AspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320
Qy      1498 TTAGACTTCAGAGATTTATAGATCTCGTATTAAGGCTTAACTACTTACCAAACTTC 1557
Db      321 LeuAspPheArgAspLeuIleArgProArgAspLysAlaIleLeuLeuIleuValAsnLeu 340
Qy      1558 GATGCTTTTGGTATTTATGACATTTACTTAACTGAAAAGTAGAGATTAATCAGATGAC 1617
Db      341 AspAlaPheGlyIleMetAspIleThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
Qy      1618 ACCAACCGTATCAACCGTTTATATGGGCAAGCGACCGGAGAGAGAGATGTTACTTAT 1677
Db      361 ThrAsnAlaGlyIleThrValIleThrGlyLysArgProGlnGlyGluAsnAlaSerIle 380
Qy      1678 CATTAGCTATGATTAAGATCGTTTATACCGAAGAAAGACGAAATTTTACAGCTACCTG 1737
Db      381 HisLeuAlaIleArgLysAspArgIleThrGlnGlnGluValIleIleSerIleLeu 400
Qy      1738 CGTTATACAGGACACCTTATACCTGATTAACCTTAAACGACAAA 1779
Db      401 ArgIleThrGlyThrProIleProAspAsnProAsnAspLys 414

```

RESULT 5
US-08-560-098A-52

```

? Sequence 52, Application US/08560098A
? Patent No. 5976841
? GENERAL INFORMATION:
? APPLICANT: INENDT, Stephan
? APPLICANT: HEINZEL-WIELAND, Regina
? APPLICANT: STEFFENS, Gerd Josef
? TITLE OF INVENTION: Proteins having fibrinolytic and
? TITLE OF INVENTION: Coagulation-inhibiting Properties
? NUMBER OF SEQUENCES: 60
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
? STREET: 1200 G Street, N.W., Suite 700
? CITY: Washington
? STATE: DC
? COUNTRY: USA
? ZIP: 20005
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent In Release #1.0, Version #1.30 (ERO)
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/560,098A
? FILING DATE: 17-NOV-1995
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: P 44 40 892.7
? FILING DATE: 17-NOV-1994
? ATTORNEY/AGENT INFORMATION:
? NAME: EVANS, Joseph D.
? REGISTRATION NUMBER: 26,269
? REFERENCE/DOCKET NUMBER: 148/42448
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (202) 628-8800
? TELEFAX: (202) 628-8844
? INFORMATION FOR SEO ID NO: 52:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 440 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-08-560-098A-52

Alignment Scores:
Pred. No.: 3,99e-210 Length: 440
Score: 2158.00 Matches: 414
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 68.86% Indels: 0
Gaps: 0
DB: 1

US-09-940-235-11 (1-1782) x US-08-560-098A-52 (1-440)
Qy      538 ATTGCTGACGACCTGATGCGCTAGACCGTCCATCTGTCAACACAGCAATTGGTGT 597
Db      27 LLeAlaGlyProGluIleuPheLeuAspArgProSerValAsnSerGlnLeuValVal 46
Qy      598 AGCGTTCGTGTCTGTGTTGAGGGGACAAATCAAGACATTAGCTTAAATTTTGAATC 657
Db      47 SerValAlaGlyThrValGlnGlyThrAsnGlnAspIleSerLeuLysPheGluIle 66
Qy      658 GATCTAACATCACGACCTGCTCATGAGGAAGAAGACAGACGAAAGGCTTAAGTCCAAATCA 717
Db      67 AspLeuThrSerArgProAlaHisGlyGlyThrGlnGlnGluIleuSerProLysSer 86
Qy      718 AAACCATTTGCTACTGATGATGCGCGGATGTCACATTAACCTTGAAGAAAGCTGACTTACTA 777
Db      87 LysProPheAlaThrAspSerGlyAlaMetSerHisIleuLeuGluValAspLeuLeu 106
Qy      778 AAGGCTATTCAAGAACATTTGATCGCTTACGTCACAGTAACGACGATCTTTGAGTGC 837
Db      107 LysAlaIleGlnGlnGlnIleuIleAlaAsnValHisSerAsnAspIlePheGluVal 126

```



```

QY 838 ATGATTTTGCAGGAGTGAACAACATTAATGATCGAAGGCGAAGGCTTACTTGGCTGAC 897
Db 127 ILeaspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLeValGlyrheAlaAsp 146
QY 898 AAAAGATGTTGGTAACTTGGCCGACCCAGCTGTGCAAGAAATTTTGGTAAAGCGGACAT 957
Db 147 LysAspGlySerValThrLeuProThrGlnProValGlnGlnPheLeuSerGlyHis 166
QY 958 GTGGCGGTTAGACATATATAAGAAAAACAATATACAAAACAAGCGAAATCTGTGATGG 1017
Db 167 ValArgValArgProGlyrLeuSgIuysProIleGlnAsnGlnAlaLysSerValAspVal 186
QY 1018 GAATATATCTGTACAGTTTACTCCCTTAAACCTGTATGACGATTTGACAGAGCTTCAAA 1077
Db 187 GluTyThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 206
QY 1078 GATACATAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCAATCTCAAGAAATTA 1137
Db 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeuLeu 226
QY 1138 GCTCAGACCAAGAAGATTTTAAACAAAACAAGCGGTATACGATTTATGAACTGAC 1197
Db 227 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyThrIleTyGlnArgAsp 246
QY 1198 TCCTCAATCGTCACTCATGACATGACATTTTCCGTAAGATTTTACCAATGATCAAGAG 1257
Db 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGln 266
QY 1258 TTTAATTAACCGTGTAAAAAATCGGGAACAAGCTTATAGGATCAATAAAAATCTGGTCTG 1317
Db 267 PheThrTyArgValLysAsnArgGlnGlnAlaTyArgIleAsnLysSerGlyLeu 286
QY 1318 AATGAAGAAAATAACACACATGACCTGATCTGTGAAAAATATTAGCTCTTAAAAAGG 1377
Db 287 AsnGlnGlnLysLeuAsnThrAspLeuLysSerGlnLysTyTyValLeuLysGly 306
QY 1378 GAAAGCCGTAATGATCCCTTGTGACGAGTCACTGGAACCTGTACCATCAATACGTT 1437
Db 307 GlnLysProTyArgProPheAspArgSerHisLeuLysLeuPheThrIleLysTyVal 326
QY 1438 GATTCGATACCAACGAATGCTTAAAAAAGTGAAGAGCTTCTTAAAGCTACGCAACGTAAC 1497
Db 327 AspValAspThrAsnGlnLeuLeuLysSerGlnGlnLeuThrAlaSerGlnArgAsn 346
QY 1498 TTAACCTTCAAGAGATTTTATGATCTGTGATTAAGGCTTAACTTCTTACCAACATCTC 1557
Db 347 LeuAspPheArgAspLeuTyArgProArgAspLysAlaLysLeuLysTyArgAsnLeu 366
QY 1558 GATGCTTTTGGTAATATAGCATATACCTTAACTGAAAAAGTAGAGATTAATACAGATGAC 1617
Db 367 AspAlaPheGlyIleMetAspTyThrLeuThrGlyLysValGlnAspAsnHisAspAsp 386
QY 1618 ACCAACGATATCATTAACCGTTTATATGGCAAGACCCGGAAGAGAGATCTAGCTAT 1677
Db 387 ThrAsnArgIleIleThrValTyMetGlyLysArgProGlnGlnGlnAsnHisLeuLys 406
QY 1678 CATTTAGCTTATGATTAAGATCGTTTATACCGAAGAAAGAAAGATTATACGATCTG 1737
Db 407 HisLeuAlaTyArgLysAspArgTyThrGlnGlnGlnArgGlnValTyLysSerTyLeu 426
QY 1738 CGTTATACAGGAGACCTTATACCTTAACTTAACTTAACTTAACTTAACTTAACTTAA 1779
Db 427 ArgTyThrGlnGlnProIleProAspAsnProAsnAspLys 440

```

```

; CURRENT APPLICATION NUMBER: US/10/360,101
; CURRENT FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: EP 02077060.8
; PRIOR FILING DATE: 2002-05-24
; NUMBER OF SEQ ID NOS: 309
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 264
; LENGTH: 413
; TYPE: PRF
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: sequence of streptokinase
US-10-360-101-264

Alignment Scores:
Pred. No.: 1,24e-209 Length: 413
Score: 2153.00 Matches: 413
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 68.70% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-11 (1-1782) x US-10-360-101-264 (1-413)
QY 538 ATGCTGACCTGAGTGGCTGAGACCGTCAATCTGTCAACAACAGCCAAATGTTGTT 597
Db 1 ILeAlGlyProGlnTyThrLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal 20
QY 598 AGCGTGTCTGTACTGTGTGAGGAGCAAGATCAAGACATTAAGTCTTAAATTTTGAATC 657
Db 21 SerValAlaGlyThrValGlnGlnTyThrAsnGlnAspIleSerLeuLysPhePheGlnIle 40
QY 658 GATCTTAACATCACGACCTGTCTGAGAGAAAGACAGACGAAAGCTTAAAGTCAAAATCA 717
Db 41 AspLeuThrSerArgProAlaHisGlyGlyLysThrGlnGlnGlnLysSerProLysSer 60
QY 718 AAAACATTTGCTATGATATGAGGAGGATGTCATATACTTGAAAGCGATTACTA 777
Db 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlnLysAlaAspLeuLeu 80
QY 778 AAGGCTATTCAGAACAAATGATCGCTAACGTCACAGTACAGACGATCTTGAAGTGC 837
Db 81 LysAlaIleGlnGlnGlnLeuIleAlaSerValHisSerAsnAspArgTyrPheGlnVal 100
QY 838 ATGATTTTGCAGGAGTGAACCATTAATCTGATGAAACGCAAGGCTTACTTGTCTGAC 897
Db 101 ILeAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyPheAlaAsp 120
QY 898 AAAAGATGTTGGTAACTTGGCCGACCCAGCTGTGCAAGAAATTTTGGTAAAGCGGACAT 957
Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnGlnPheLeuSerGlyHis 140
QY 958 GTGGCGGTTAGACATATATAAGAAAAACAATACAAAACAAGCGAAATCTGTGATGATG 1017
Db 141 ValArgValArgProTyArgLysGlnLysProIleGlnAsnGlnAlaLysSerValAspVal 160
QY 1018 GAATATATCTGTACAGTTTACTCCCTTAAACCTGTATGACGATTTGACAGAGCTTCAAA 1077
Db 161 GluTyThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
QY 1078 GATACATAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCAATCTCAAGAAATTA 1137
Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeuLeu 200
QY 1138 GCTCAGACCAAGAAGATTTTAAACAAAACAAGCGGTATACGATTTTATGAACTGAC 1197
Db 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyThrIleTyGlnArgAsp 220
QY 1198 TCCTCAATCGTCACTCATGACATGACATTTTCCGTAAGATTTTACCAATGATCAAGAG 1257
Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGln 240
QY 1258 TTTAATTAACCGTGTAAAAAATCGGGAACAAGCTTATAGGATCAATAAAAATCTGGTCTG 1317

```

Db 241 PheThrTyrArgValIlysaAnaArgGluGlnAlaIlyrArgIleAenIlySvsSerGlyLeu 260
QY 1318 AATGAAGAATAAACAACACTGACCTGATCTCTGAGAAATATTACGCTCTTAAAAAAGGG 1377
Db 261 AsnGluGluIleAsnAnthrAspLeuIleSerGluIlySlyTyrValIleuIlySlyGly 280
QY 1378 GAAAGCCGTATGATCCCTTGTATCGCAGTCACTTGAACCTGTTCAACATCAAAATACGTT 1437
Db 281 GluIlyProTyrAspProPheAspArgSerHisIleuIlySleuPheThrIleIySlyrVal 300
QY 1438 GATGTGATACCAACGAATTCGTAAAAAGAGAGAGCTCTTAACAGCTAGCGAAACGTAC 1497
Db 301 AspValAspThrAsnGluIleuIlySvsSerGluGlnIleuIlyrAlaIleSerGluArgAn 320
QY 1498 TTGACCTTGAGAGATTTATACGATCCCTCGATTAAGGCTAAACACTCAACAAACTC 1557
Db 321 LeuAspPheArgAspLeuIlyrAspProArgAspIlyAlaIlySleuIlyrAsnAsnIleu 340
QY 1558 GATGCTTTGGTATATATGACTATACCTTAACCTGAAAGAGATGAATATCAGATGAC 1617
Db 341 AspAlaPheGlyIleMetAspTyrThrIleThrGlySlyValGluAspAsnHisAspAsp 360
QY 1618 ACCGACCGTATCAATACCGTTTATATGCGCAAGCGACCCGAAAGAGAGATGCTAGCTAT 1677
Db 361 ThrAsnArgIleIleThrValIlyrMetGlyIlyArgProGluGluIlyAlaSerTyr 380
QY 1678 CATTTAGCCATATGATTAAGATCGTTTATACGGAAGAAAGAAAGATTTACAGTACTG 1737
Db 381 HisLeuAlaIlyrAspIlySvsAspArgTyrThrGluGluIlyrGluValIlyrSerTyrIleu 400
QY 1738 CGTTATACAGGACACCTTATACCTGATTAACCTTAACGAC 1776
Db 401 ArgTyrThrGlyThrProIleProAsnProAsnAsp 413

RESULT 7
US-08-488-940-1
Sequence 1, Application US/08488940
Patent No. 5854049
GENERAL INFORMATION:
APPLICANT: Reed, Guy L.
TITLE OF INVENTION: PLASMIN-RESISTANT STREPTOKINASE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSER: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,940
FILING DATE: 09-JUN-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 34,819
REFERENCE/DOCKET NUMBER: 05433/009001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1194 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear

MOLECULE TYPE: protein
US-08-488-940-1
Alignment Scores:
Pred. No.: 3,82e-209 Length: 1194
Score: 2150.50 Matches: 420
Percent Similarity: 97.68% Conservative: 1
Best Local Similarity: 97.45% Mismatches: 5
Query Match: 68.62% Indels: 5
DB: 1 Gaps: 3
US-09-940-235-11 (1-1782) x US-08-488-940-1 (1-1194)
QY 496 CAGACCAATGAGAGC-----GGATCTGGCCCTTACCGGATGTTGGATTGCTGGA 546
Db 366 GlnThrAsnSerSerSerValProGlyArgGlySerIle---GluGlyArgIleAlaGly 384
QY 547 CTTGAGTGGCTGTAGACCGTTCATCTGTCAACAACACCAATGGTTGTTAGCGTTGCT 606
Db 385 ProGluThrLeuIleuAspArgProSerValAsnAsnSerGlnLeuValValSerValAla 404
QY 607 GGTACTGTGGGGGAGCAATCAAGACATTAGTCTTAATTTTGAATGATCTTAACA 666
Db 405 GLyThrValGluGlyThrAsnGlnAspIleSerIleuIlyrPheGluIlyIleAspLeuThr 424
QY 667 TCAGACCGCTCATGAGAGAAAGCAGAGCAAGGCTTAAGTCAAAATCAAAACATTT 726
Db 425 SerArgProAlaHisGlyGlyIlyrThrGluGlnGlyLeuSerProIlySerIlyrPhe 444
QY 727 GCTACTGATAGTGGCGCGATGTCATTAACCTTGAGAAAGCTGACTTAAGGCTATT 786
Db 445 AlaThrAspSerGlyAlaMetSerHisIlySleuGluIlyrAlaAspLeuIlyAlaIle 464
QY 787 CAAGAACATTTGATCGCTTAAGTCCACAGTACAGACACTACTTTGAGGTCAATTGATTT 846
Db 465 GlnGluGlnIleuIleAlaAsnValHisSerAsnAspArgTyrPheGluValIleAspPhe 484
QY 847 GCAAGCGATGCAACATTAATCATGATCGAAAGCGCAAGGCTCTTTGCTGCAAAAGATGCT 906
Db 485 AlaSerAspAlaThrIleThrAspArgAsnIlyrSlyValIlyrPheAlaAspIlyrAspIly 504
QY 907 TCGGTAACTTGGCCGACCAACCTGTCCAAGAAATTTTGTAAAGCGGACATGTGGCGCTT 966
Db 505 SerValThrLeuProThrGlnProValGlnIlyrLeuLeuLeuSerGlyHisValArgVal 524
QY 967 AGACCATATTAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTGAATATACT 1026
Db 525 Arg---TyrIlySlyIlyrProIleGlnAsnGlnAlaIlyrSerValAspValGlyThr 543
QY 1027 GTACAGTTTATCCCTTAAACCTGATGACGATTTTCAGACCAAGGCTTCAAGATATCAAG 1086
Db 544 ValGlnPheThrProLeuAsnProAspAspArgProGlyLeuIlyrAspThrIlyS 563
QY 1087 CTATTGAAACACTAGCTATGCTGAGACCACTGACATGCTCAAGAAATTTCTGTGCTCAAGCA 1146
Db 564 LeuIleuIlyrThrIleuAlaIleGlyAspThrIleThrSerGlnIlyrLeuIlyrGlnAla 583
QY 1147 CAAGCATTTTAAACAAACCAACCAAGGCTATACGATTTTATGACGTAACCTCAATC 1206
Db 584 GlnSerIleIleuAsnIlyrAsnHisProGlyTyrThrIleIyrgluIlyrAspSerSerIle 603
QY 1207 GTCACTCATGACATGACATTTTCCGTACGATTTTACCATGATGAAGATTTACTTAC 1266
Db 604 ValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGluIlyrPheThrTyr 623
QY 1267 CGTGTAAAAATCGGGAACAAGCTTATAGATTCATTAATAAATCTGGCTGAAGAA 1326
Db 624 ArgValIlyrAsnAspArgGluGlnAlaIlyrArgIleAsnIlyrSvsSerGlyLeuAsnGluIly 643
QY 1327 ATAAACAACACTGATCTGTGAGAAATATTACGCTTAAAAAAAGGGAAGACCG 1386
Db 644 IleAsnAnthrAspLeuIleSerGluIlySlyTyrValIleuIlyrSlyGlyGluIlyrPro 663

Qy	1387	TATGATGCCCTTGAATCGGAGCACTTGAAACGTTCACCATCAATACGTGAGTGCAT	1446
Db	664	TyrAspProPheAspArgSerHisIleuysIeuPheTrIleYsYrValAspValAsp	683
Qy	1447	ACCAACGAATTCCTAAAAGTAGAGCGCTCTTAAACAGCTAGCGAAGCTTAACCTTACACTTC	1506
Db	684	ThrAnGIuIeuIeuIeuYSerGIuGIuIeuIeuThrIaSerGIuIaGAsIeuAspHe	703
Qy	1507	AGAGATTATACGATCCTCGTGATTAAGCTTAACCTACTCTACAACAATCTCGATGCTTTT	1566
Db	704	ArgAspIeuYrAspProArgAspIysAlaIysIeuIeuYrAsnAsnIeuAspAlaPhe	723
Qy	1567	GGTATTATAGCACTTATACCTTAACCTGAGAAAGTAGAGATATATCAAGATGACCAACCGT	1626
Db	724	GlyIleThrValIyrThrIeuThrGIyysValGIuAspAsnHisAspAspThrAsnArg	743
Qy	1627	ATCATTAACCGTTTATATGAGCAAGCAACCCGAGAGAGAGATGCTAGCTATCATTTAGCC	1686
Db	744	IleIleThrValIyrMetGIyIysArgProIuGIuIaAsnAlaSerYrHisIeuAla	763
Qy	1687	TATGATTAAGATCGTTATATACCGAAGAGAAGAGAAAGATTACAGCTACCTGCTTATACA	1746
Db	764	TyrAspIysAspArgYrThrGIuGIuIaArgIuValIyrSerYrIeuArgYrThr	783
Qy	1747	GGGACACCTATACCTGATTAACCTTAACGACAAA	1779
Db	784	GlyThrProIleProAspAsnProAsnAspIys	794
RESULT 8			
US-07-854-596B-28			
Sequence 28, Application US/07854596B			
Patent No. 5434073			
GENERAL INFORMATION:			
APPLICANT: Dawson, Keith M			
APPLICANT: Hunter, Michael G			
APPLICANT: Czaplowski, Lloyd G			
TITLE OF INVENTION: Proteins and nucleic acids			
NUMBER OF SEQUENCES: 73			
CORRESPONDENCE ADDRESS:			
ADDRESSER: Dr. John J. McDonnell			
STREET: Ten South Wacker Drive, Suite 3000			
CITY: Chicago			
STATE: IL			
COUNTRY: USA			
ZIP: 60606			
COMPUTER READABLE FORM:			
MEDIUM TYPE: Floppy disk			
COMPUTER: IBM PC compatible			
OPERATING SYSTEM: PC-DOS/MS-DOS			
SOFTWARE: PatentIn Release #1.0, Version #1.25			
CURRENT APPLICATION DATA:			
APPLICATION NUMBER: US/07/854,596B			
FILING DATE: 03-JUN-1992			
CLASSIFICATION: 435			
ATTORNEY/AGENT INFORMATION:			
NAME: McDonnell, John J			
REGISTRATION NUMBER: 26,949			
REFERENCE/DOCKET NUMBER: 92,337			
TELECOMMUNICATION INFORMATION:			
TELEPHONE: 312-715-1000			
TELEFAX: 312-715-1234			
TELEX: 910-221-5317			
INFORMATION FOR SEQ ID NO: 28:			
SEQUENCE CHARACTERISTICS:			
LENGTH: 499 amino acids			
TYPE: amino acid			
TOPOLOGY: linear			
MOLECULE TYPE: protein			
US-07-854-596B-28			
Alignment Scores:			
Pred. No.: 1.42e-208			
Score: 2143.00			
Length: 499			
Matches: 413			

Percent Similarity:	96.29%	Conservative:	2
Best Local Similarity:	95.82%	Mismatches:	16
Query Match:	68.38%	Indels:	0
DB:	1	Gaps:	0

US-09-940-235-11 (1-1782) x US-07-854-596B-28 (1-499)

QY	487	ACCTCTGCGACGACCATCGAGGGGATCTGCGCCCTTCAACCATGTTCCGATTTGCTGGA	546
Db	69	ThrlleaserllealaalalysgluglglValserleuaplysarglIealagl	88
QY	547	CCCTAGAGGCGCTGACCGGTCATCTGTCAACAAAGCCAAATGGTTGGTAGGGTGGCT	606
Db	89	ProclutrlpeuleuwaspargProSeValasmnserglneulValSeVala	108
QY	607	GGTACTGTGGAGGAGCAATCAAGCATTTAGTCTTAAATTTTGGAAATCGATCTACA	666
Db	109	glYthralgluglYthraenglnaprlIeserleuysphepneglulleapneuthr	128
QY	667	TCACGACCTGCTCATGGAAGAAAGACAGAGCGCTTAAGTCCAAATCAAAACATTT	726
Db	129	SezarprProclanhsrglYglYlYsthluglglneulYleuserProlysSerlYsProphe	148
QY	727	GCTACTGATGTGGCGGATGTGCATPAACTTGAGAAAGTGACTTAATCTPAAAGGCTATT	786
Db	149	AlathrasperglYalameProhlsleuulYulysalaepleuYalale	168
QY	787	CANAGCAAAATTGATCGGTCACGTCACAGTAAACAGCATCTTGTGAGTCATTGATTTT	846
Db	169	GlnglglneulIealaenValHlsSerlasmnspargYrphneglVallleasphe	188
QY	847	GCAAGCATGCAACCATTTACTGATGAAAGCGCAAGGCTTACTTTTGCTGACAAAGATGT	906
Db	189	AlaseraplatrlthrlIethrasparYasnglYysValYrphelIaasplyaspglY	208
QY	907	TGCGTAACTTTGCCGACCCACTGTCGCAAGAAATTTTGTCAACGGACATGTCGGCTT	966
Db	209	SerValThrlneupProthrlnProValGlnglneuleuSerglYHlsvalArYal	228
QY	967	AGACCATPATAAGAAAAACCAATCAAAACCAAGCAAACTCTGTGATGTCGATATPACT	1026
Db	229	ArgProYrlyrsbgulysProllleglnasnlnalalysSerValIasplValglutYrThr	248
QY	1027	GTACAGTTTACTCCCTTAAACCTGATGACGATTTTGACACAGGCTCAAGATPACTAG	1086
Db	249	ValGlnpethrProleuasnProasparspheargProglYleuYasprThrlys	268
QY	1087	CTATTGAAAAACCTAGCTATCGGTGACACCATCATCTCAAGAAATTCTAGCTCAAGCA	1146
Db	269	leuYeuYsThrlleuallleglYasprThrIethrSerlnglneuleuValleaglala	288
QY	1147	CAAGCATTTTAAACAAAACCAACCCGAGGCTATCGAATTAATGAAACGATCTCCCTCATC	1206
Db	289	GlmserrlleuasnlystnhrIsproglYrYrThrIleYrglnaYasprSerlle	308
QY	1207	GTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGATCAAGATTTACTTAC	1266
Db	309	ValtrhrIsapnaenaprlIephargThrIleuProhetasprnglInpethrTyr	328
QY	1267	CGTGTAAAAATCGGGAACAACTTATAGGATCAATAAAAATTTGGTCTGAATGAAGAA	1326
Db	329	HlsVallysnlnalYglnglnalalYrglnlleasnlylsSerglYleuasnlnlglu	348
QY	1327	ATPAAACAAACCTGACCTGATCTGCAAAAATATTAAGCTCTTAAAAAGGGGAAAAGCGC	1386
Db	349	IlasmnsthraapleullIeserglulysYrYrValleuYslysglYulnYsPro	368
QY	1387	TATGATCCCTTGTATCGCAGTCACTTGAAATCTGTTCACATCAAAATCGTTGATGTGCAT	1446
Db	369	TyrAsprProphesapargSerHlsleuYlsleuphethrIleuYsYrValIasprValasn	388
QY	1447	ACCAAGCAATTTGTAATAAGTAGACAGCTTTAACAGTACGCAAGTAACTTACATTC	1506

Db 389 ThrAnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAnLeuLeuAspPhe 408
 QY 1507 AGAGATTATACGATCTCGTGTAGTAAGCTAACTACTACAAACAATCTCGATGCTTTT 1566
 Db 409 ArgAspLeuTyrAspProArgAspLysAlaIleLysLeuTyrAsnAnLeuAspAlaPhe 428
 QY 1567 GGTATTATGACATATACCTTAACCTGAGAAAGTAGAGATATACACGATGACACCAACCGT 1626
 Db 429 GlyIleMetAspTyrThrLeuThrGlyLysValGluAspAnHisAspAspThrAsnArg 448
 QY 1627 ATCATTAACCGTTTATATGCGCAAGCACCAGAGAGAGAAATGCTACTTCACTTTAGCC 1686
 Db 449 IleIleThrValTyrMetGlyLysArgProGluGlnLysAnHisAspTyrHisLeuAla 468
 QY 1687 TATGATAAAGATCGTTATACCGAAGAAAGAGAGATTACAGCTACCTGCGTTATACA 1746
 Db 469 TyrAspLysAspArgTyrThrGlnGluGlnArgGlnValTyrSerTyrLeuArgTyrThr 488
 QY 1747 GGGACACCTATACCTGATTAACCTTAACGACAAA 1779
 Db 489 GlyThrProIleProAspAnProAsnAspLys 499

RESULT 9

US-08-488-940-3
 ; Sequence 3, Application US/08488940
 ; Patent No. 5854049

; GENERAL INFORMATION:
 ; APPLICANT: Reed, Guy L.
 ; TITLE OF INVENTION: PLASMIN-RESISTANT STREPTOKINASE
 ; NUMBER OF SEQUENCES: 20
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fish & Richardson P.C.
 ; STREET: 225 Franklin Street
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02110-2804

; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/488,940
 ; FILING DATE: 09-JUN-1995
 ; CLASSIFICATION: 514
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Frazer, Janis K.
 ; REGISTRATION NUMBER: 34,819
 ; REFERENCE/DOCKET NUMBER: 05433/009001
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 617/542-5070
 ; TELEFAX: 617/542-8906
 ; TELEX: 200154
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 813 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: not relevant
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-488-940-3

Alignment Scores:
 Pred. No.: 3,266-208 Length: 813
 Score: 2140.50 Matches: 413
 Percent Similarity: 99.76% Conservative: 0
 Best Local Similarity: 99.76% Mismatches: 0
 Query Match: 68.30% Indels: 1
 Gaps: 1

US-09-940-235-11 (1-1782) x US-08-488-940-3 (1-813)

QY 538 ATTGCTGACCTGATGCGCTCTAGACCGTCCATCTGTCAACACAGCAATGCTGTT 597
 Db 1 IleIleGlyProGluTutPrLeuLeuAspArgProSerValAsnAspSerGlnLeuVal 20
 QY 598 AGCGTGTGCTACTGTTGAGGGGACGAATACAGCATTTAGTCTTAAATTTTGAATC 657
 Db 21 SerValIleGlyThrValGlnGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 40
 QY 658 GATTCATCATGACGACCGCTCATGAGGAGAAAGACAGGCAAGGCTTAAGTCCAAATCA 717
 Db 41 AspLeuThrSerArgProAlaHisGlyGlyLysThrGlnGlnGlyLeuSerProLysSer 60
 QY 718 AACCAATTTGCTACTGATAGTGGCGGATGCACTAATACTTGAGAAAGCTGACTACTA 777
 Db 61 LysProPheAlaThrAspSerCylAlaMetSerHisLysLeuGluLysAlaAspLeu 80
 QY 778 AAGCTATTCAGAAACAATTGATGCTTACCTCCACAGTAAACAGACGACTTTGAGGTC 837
 Db 81 LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100
 QY 838 ATTGATTTCGAAACGATGCAACCATTAATGATGGAACGGCAGAGTCTACTTGTCTAC 897
 Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
 QY 898 AAAGATGATTCGGTAACCTTGCCGACCCCAACCTGTCCAAAGAAATTTTGTAAAGCGACAT 957
 Db 121 LysAspLysSerValThrLeuProThrGlnProValGlnGlnPheLeuLeuSerGlyHis 140
 QY 958 GTGGCGGTAGACCATTAATAAGAAACCAATACAAACCAAGCAAAATCTGTGATGTG 1017
 Db 141 ValArgValArg--TyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 159
 QY 1018 GAATATCTGATACGTTTACTTCCTTAAACCTGATGACGATTTCAGACCGGTCTCAA 1077
 Db 160 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 179
 QY 1078 GATCTAGGCTATGAAAACACTAGCTATGCGTGAACCATCACTCAAGAAATTTCTA 1137
 Db 180 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnLeuLeu 199
 QY 1138 GCTCAACGACAAAGCATTTTAAACAAACCAACCCAGGCTTATGACATTTATGACGTGAC 1197
 Db 200 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 219
 QY 1198 TCCTCAATCGTCACTGACATGACATGACATTTCCGTACGATTTTCAAGATGACAG 1257
 Db 220 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 239
 QY 1258 TTTACTTACCGTGTAAATAATCGGGAACAGCTTATGATCATTAATAATAATCGGTCTG 1317
 Db 240 PheThrTyrArgValLysAsnArgGlnGlnAlaTyrArgIleAsnLysSerGlyLeu 259
 QY 1318 AATGAAGAAATTAACAACACTGACCTGATCTGAGAAATATTAAGTCTTAAAAAGGG 1377
 Db 260 AsnGluGluIleAsnAnHisAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 279
 QY 1378 GAAAAGCGTATGATCCCTTGTATCGAGTCACTTGAACCTGTTACCATCAAAATAGCTT 1437
 Db 280 GlyLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 299
 QY 1438 GATGTCGATACCAACGAATGCTAATAAGTAGAGAGCTTTTAAACGCTAGGAAAGTAC 1497
 Db 300 AspValAspThrAsnGluLeuLeuLysSerCylGlnLeuLeuThrAlaSerGluArgAsn 319
 QY 1498 TTAGACTTCAGAGATTATATCGANTCCGCTGATTAAGGCTAAACATCTCAACAATCTC 1557
 Db 320 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 339
 QY 1558 GATGCTTTGTATATGACTATACCTTAAGTGAAGAAAGTAGAGATTAATCAAGATGAC 1617
 Db 340 AspAlaPheGlyIleMetAspTyrThrLeuThrCylLysValGluAspAnHisAspAsp 359
 QY 1618 ACCAACCGTATCATTAACGTTTATATGAGGCAAGGACCCGAAGAGAAATGCTAGCTAT 1677

```

Db      360  ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGluLeuAlaSerTyr 379
Qy      1678  CATTAGCCTTAGTAAGAATGCTTATACCGAAGAAGACGAACTTACCTACTG 1737
Db      380  HisLeuAlaTyrAspLysAspArgTyrThrGluGluArgGluValTyrSerTyrLeu 399
Qy      1738  CGTTATACAGGAGACCTTATACCTGATACCTTACCGAAGAA 1779
Db      400  ArgTyrThrGlyThrProIleProAspAsnProAsnAspLys 413

RESULT 10
US-07-854-596B-26
; Sequence 26, Application US/07854596B
; Patent No. 5434073
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith M
; APPLICANT: Hunter, Michael G
; APPLICANT: Czaplowski, Lloyd G
; TITLE OF INVENTION: Proteins and nucleic acids
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Dr. John J. McDonnell
; STREET: Ten South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/854,596B
; FILING DATE: 03-JUN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26, 949
; REFERENCE/DOCKET NUMBER: 92,337
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 415 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-854-596B-26

Alignment Scores:
Pred. No.: 1,696-207 Length: 415
Score: 2132.00 Matches: 409
Percent Similarity: 99.03% Conservative: 1
Best Local Similarity: 98.79% Mismatches: 4
Query Match: 68.03% Indels: 0
Gaps: 0

US-09-940-235-11 (1-1782) x US-07-854-596B-26 (1-415)
Qy      538  ATTGCTGACCTGATGAGCTGCTAGACCTGCATCTGTCAACAACGCAATTGCTTGT 597
Db      2  IleAlaGlyProGluTyrLeuLeuAspArgProSerValAlaAsnSerGlnLeuVal 21
Qy      598  AGCGTGTCTGCTACTGTTGAGGGAGCAATCAAGACATTAGCTTAAATTTTGAATC 657
Db      22  SerValAlaGlyThrValGluGluGlyThrAsnGlnAspIleSerLeuYsPheGluIle 41
Qy      658  GATCAACATACGACGCTGCTCATGAGAGAAAGACGACGACGCTTAAGTCCAAATCA 717

```

```

Db      42  AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 61
Qy      718  AAACCATTTGCTATGATAGTGGCGCATGTCACATTAACCTTGAGAAAGCTGACTACTA 777
Db      62  LysProHeaIatThrAspSerGlyAlaMetProHisLysLeuGluValAspLeuLeu 81
Qy      778  AAGCTATTCCAAAGAACATTTGATGCTTACGCTCCACAGTAACGACGACTTGTGAGGCT 837
Db      82  LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 101
Qy      838  ATTGATTTTGGACGATGACGATGACCATTAATGATGCGAAACGGCAAGTCTTGTGTCAGC 897
Db      102  IleAspPheAlaSerAspAlaThrIleThrAspArgGlnGlyValTyrPheAlaAsp 121
Qy      898  AAAGATGCTTCGGTACCTTGGCCGACCCCAACCTGTCCAAAGAAATTTTGTCTAAGCGGACAT 957
Db      122  LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 141
Qy      958  GTGCGCGTTAGACCATTAAGAAAGAAACCAATACAAAGCAAGGAAATCTGTGATGTG 1017
Db      142  ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 161
Qy      1018  GAATATACGTGACGATTACTCCCTTAAACCTGATGACGATTTCCAGACGATCTCAAA 1077
Db      162  GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 181
Qy      1078  GATATCAAGCTATTGAAACACATGCTATCGGTGACACCATCACTCAAGATTACTA 1137
Db      182  AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnLeuLeu 201
Qy      1138  GCTCAACGACAAAGCATTTTAAACAAAGAACACCGAGCTTACGATTGAAAGCTGAC 1197
Db      202  AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp 221
Qy      1198  TCCTCAATGCTCATCTGACATGACATGATTTCCGTACGATTTTACCAATGATCAAG 1257
Db      222  SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetCAspGlnGlu 241
Qy      1258  TTATCTTACCGTGTAAATAATCGGAAACAGCTTATAGATCAATPAAAAATCTGCTCTG 1317
Db      242  PheThrTyrHisValLysValLysAsnArgGluGlnAlaTyrGluIleAsnLysLysSerGlyLeu 261
Qy      1318  AATGAAAGAAATPAAACACATGACCTGATCTCTGAGAAATTTACGCTTAAAAAAGG 1377
Db      262  AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysGlyLeu 281
Qy      1378  GAAAAGCCGTATGATCCCTTTGATCGGAGTCACTTGAACGTTCACCATCAAAATACGTT 1437
Db      282  GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 301
Qy      1438  GATGTCGATACCAACGAATTGCTTAAAAAGTAGAGCAGCTTTAACAGCTAGCGAACGTAAAC 1497
Db      302  AspValAsnThrAsnGlnLeuLeuLysSerGlnGlnLeuLeuThrAlaSerGluArgAsn 321
Qy      1498  TTAGACTTCAGAGATTTTATACGATCCCTCGTATAGGATTAAGCTTAACTTACAAATCTC 1557
Db      322  LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 341
Qy      1558  GATGCTTTTGGTATTATGACTATACCTTAACTGGAAGAAAGTAGAGATTAACGATGAC 1617
Db      342  AspAlaPheGlyIleMetLysPyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 361
Qy      1618  ACCAACGCTATCATTAACCGTTTATATGGGACAGGACCCGGAAGAGAGAATGCTAGCTAT 1677
Db      362  ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGluGluAlaAsnAlaSerTyr 381
Qy      1678  CATTACCTTATGATTAAGAATCGTTATACCGAAGAAGAAAGAAAGCTTACGCTACTG 1737
Db      382  HisLeuAlaTyrAspLysAspArgTyrThrGluGluGluValTyrSerTyrLeu 401
Qy      1738  CGTTATACAGGAGACCTTATACCTGATTAACCTTACGACGAA 1779
Db      402  ArgTyrThrGlyThrProIleProAspAsnProAsnAspLys 415

```

RESULT 11
US-07-854-596B-19
; Sequence 19, Application US/07854596B
; Patent No. 5434073
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith M
; APPLICANT: Hunter, Michael G
; APPLICANT: Czaplewski, Lloyd G
; TITLE OF INVENTION: Proteins and nucleic acids
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. John J. McDonnell
; STREET: Ten South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/854,596B
; FILING DATE: 03-JUN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,337
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 435 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-854-596B-19

Alignment Scores:
Pred. No.: 1.73e-207 Length: 435
Score: 2132.00 Matches: 409
Percent Similarity: 99.03% Conservative: 1
Best Local Similarity: 98.79% Mismatches: 4
Query Match: 68.03% Indels: 0
DB: 1 Gaps: 0

US-09-940-235-11 (1-1782) x US-07-854-596B-19 (1-435)

QY 538 ATTGCTGACCTGATGCTGCTAGACGCTCATCTGTCAACAACGACCAATTGTTGTT 597
DB 22 Ileaaglyprrgclutrrpleuauapargprosevalashmnserrglneuvalval 41
QY 598 AGCGTGTGCTACTGTTGAGGGGAGCAATCAAGACATTAGTCTTAATTTTGAATC 657
DB 42 SerValaIaaglythrValgluglythrAanglnAaprlleSerleuysPhepegluile 61
QY 658 GATCTAACATCGACGACCTCTCTCATGGAGAAAGACAGACAGAGCTTAAGTCCAAATCA 717
DB 62 AspLeuutrSerArgrProalshIaglylysthrGluglncllyleuSerProlySer 81
QY 718 AAACCATTTGCTACTGATGAGCGCGGATGTCAACAACTTGAAGAAGCTGACTTACTA 777
DB 82 LysPrrPrrlealathrAspsergllyalawetProhlslyleugllylsalAspleuLeu 101
QY 778 AAGGCTATTCAGAACAAATTTGATCGCTAACGTCACAGATTAAGACGACTTGAAGTC 837
DB 102 LysalaIleaglInglInleuIlealAaenValHisSerAshpAspPtyrPhegluVal 121

QY 838 ATTGATTTTGCAGACGATGCAACCACTTACTGATCGAAACGGCAAGTCTACTTTGCTGAC 897
DB 122 IleaPhealAsSerApAlaThrIlethrAspArgAanglylyValTyrPhehlaAsp 141
QY 898 AAAAGATGGTTGGTAACTTCCGCAACCCAGCTGTCCAAAGAAATTTTCTTAAGCGGACAT 957
DB 142 LysAspglYserValthrLeuprOthrclnProValglngluPheleuSerclYhls 161
QY 958 GTGGCGCTTGGACATTTAAAGAAACCAATACAAACCAAGCAAGAAATCTGTGAATGTCG 1017
DB 162 ValArgValAargProTyrLysgluysProIleglInanglnAlaLysSerValAspVal 181
QY 1018 GAATATACCTGATACGTTTACTCTCCCTTAAACCCCTGATGACGATTTTCAGACCGTCTCAA 1077
DB 182 GluTyrThrValglInpethrProleuashnProAspAspPheargProglYleuys 201
QY 1078 GATATCTAGCTATTGAAAACCTAGCTATCGGTGACACCATTCACATCTCAAGAAATTACTA 1137
DB 202 AspThrlySleuLeuLysThrLeuAlaIleglYAspThrIlethrSerGlInleuLeu 221
QY 1138 GCTCAAGCACAAAGCAATTTTAAACAAAACCAACCGCTTATGATTTTGAACGTGAC 1197
DB 222 AlaGlInlaaglnSerIleleuashnLysThrshpProglYrThrIleTyrGluaArgp 241
QY 1198 TCCTCAATCGTCACTCATGACATGACATTTTCCGTGACATTTTACCAATGATCAAGAG 1257
DB 242 SerSerIleValThrshAspAsnAspIlePheargThrIleleProMetAspGlInu 261
QY 1258 TTTACTTACCGGTGTTAAAAATCGGAAACAAGCTTATAGATCAATAAAAATCTGCTCG 1317
DB 262 PheThrTyrThrshValLysAsnArggluglnAlaTyrgluileAamLysSerclYleu 281
QY 1318 AATGAAAGAAATTAACAACATGACCTGATCTCTGAGAAATTTACCTCTTAAAAAGGG 1377
DB 282 AsnGluglnIleAsnAenThrAspLeuIleSerclYusTyrTyrValLeuLysleglY 301
QY 1378 GAAAAGCGTATGATCCCTTATGTCGAGTCACTTGAACCTTCAACATCAATTAAGTT 1437
DB 302 GluysProTyrAspProPheAspArgSerhlsleuLysleuPethnIleLysTyrVal 321
QY 1438 GATGTCGATCCAAACGAATGCTTAAAAAGTAGAGGCTCTTAACAGTAGGCAAGTAAAC 1497
DB 322 AspValashnThrAanglnleuLeuLysSerclInleuLeuThrAlaSerclYuaArgshn 341
QY 1498 TTAGACTTCAGAGATTTATACGATCTCTGTGATTAAGGCTTAACTTCAACAATATCTC 1557
DB 342 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysleuLysTyrAsnAsnleu 361
QY 1558 GATGCTTTTGCTATTATGACCTTAACTTAACTGGAAGAAAGTAAACAGATAC 1617
DB 362 AspAlaPheglYIleMetAspTyrThrLeuThrglYysValGluaAspAsnshAspAsp 381
QY 1618 ACCAAGCGTATCAATTAACGTTTATATAGGCAAGGCAACCGCAAGAGAGAAATGCTAGCTAT 1677
DB 382 ThrAsnArgIleIleThrValTyrMetglYlyAsArgProgluglylsuAsnAlaSerTyr 401
QY 1678 CATTAGCTTATGATTAAGAATCGTTATATCCGAAAGACAGAGAGTTTACGACTACTG 1737
DB 402 HisleuAlaTyrAspLysAspArgTyrThrGluglnIugluValIlySerTyrleu 421
QY 1738 CGTTATACAGGACACCTTATACCTGATTAACCTTACAGCAAA 1779
DB 422 ArgTyrThrglYThrProIleProAspAsnProAsnAspLys 435

RESULT 12
US-07-854-596B-15
; Sequence 15, Application US/07854596B
; Patent No. 5434073
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith M
; APPLICANT: Hunter, Michael G
; APPLICANT: Czaplewski, Lloyd G
; TITLE OF INVENTION: Proteins and nucleic acids

```

; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. John J. McDonnell
; STREET: Ten South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/854,596B
; FILING DATE: 03-JUN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,337
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 440 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-854-596B-15

Alignment Scores:
Pred. No.: 1.74e-207 Length: 440
Score: 2132.00 Matches: 409
Percent Similarity: 99.03% Conservative: 1
Best Local Similarity: 98.79% Mismatches: 4
Query Match: 68.03% Indels: 0
DB: 1 Gaps: 0

US-09-940-235-11 (1-1782) x US-07-854-596B-15 (1-440)
QY 538 ATTGCTGACCTGAGTGGTCTAGACCCCTCATCTGTCAACAACAGCCAAATTGGTTGTT 597
DB 27 ILeaIaGlyProGluTrpLeuLeuAspArgProSeValAsnAsnSerGlnLeuVal 46
QY 598 AGCGTTGCTGTAAGTGTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTTTGAATC 657
DB 47 SerValaIaGlyThrValGlnGlnIleuValIleuValIleuValIleuValIleuVal 66
QY 658 GATTTACATCAACGACCTGCTCATGAGGAGAAAGACAGACAGGCTTAAGTCCAAATCA 717
DB 67 AspLeuThrSerArgProAlaHisGlyGlySerThrGlnGlnIleuSerProLysSer 86
QY 718 AAACATTGCTACTGATGATGAGCGCGGATGTCACATTAACCTTGAAAGCTGACTACTA 777
DB 87 LysSerProPheAlaThrAspSerGlyAlaMetProHisLysLeuGlnIleuLysAlaAspLeu 106
QY 778 AAGGCTATTCAAGAACATTTGATCGCTACGTCACAGTAAGACGACTACTTTGAGTTC 837
DB 107 LysAlaIleGlnGlnIleuValIleuValIleuValIleuValIleuValIleuValIleu 126
QY 838 ATTGAATTTGACAGGATGACCACTTACTGATCGAAGCGGACAGGCTTACTTGTGTCAC 897
DB 127 ILeaSpPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValIleuValIleuVal 146
QY 898 AAAAGTGTGGTGAACCTTCCGACCCCACTGTCGAAGAATTTTGTGAAGCGGACAT 957
DB 147 LysAspGlySerValIleuThrLeuProThrGlnProValGlnIleuPheLeuSerGlyHis 166
QY 958 GTGGCGGTTAGACCATTAATAAAGAAACCAATTAACCAAGCGGAATCTGTGATGTGG 1017

```

```

DB 167 ValArgValArgProIleuGlyGlyProIleuGlnAsnGlnAlaLysSerValAspVal 186
QY 1018 GAATATACCTGATACAGTTTACTCCCTTAAACCGTAAACGATTTCACACAGGCTCGAAA 1077
DB 187 GluTrpThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuLys 206
QY 1078 GATATACGCTATTTGAAAACATAGCTATCGGTGACACATCAATCATCTCAAGAAATTA 1137
DB 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnIleuLeu 226
QY 1138 GCTCAAGCACAAAGCATTTTAAACAAAAACACCCAGGCTATACGATTTAAGACGTGAC 1197
DB 227 AlaGlnAlaGlnSerIleuLeuAsnLysThrHisProGlyTrpThrIleuTrpGluArgAsp 246
QY 1198 TCCCTCAATGCTCACTGATGACATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 1257
DB 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
QY 1258 TTACTTACCGGTGTTAAAAATCGGAAACAGCTTATAGATCAATAAAAATCGTCTG 1317
DB 267 PheThrTrpHisValLysAsnArgGlnGlnAlaTrpGluIleuLysLysSerGlyLeu 286
QY 1318 AATGAAGAAATTAACAACACTGACTGATCTGAGAAATATTACGTCCTTAAAAAGG 1377
DB 287 AsnGlnGlnIleAsnAsnThrAspLeuIleSerGlnLysTrpTrpValIleuLysGly 306
QY 1378 GAAAAGCCGTATATATCCCTTGTATCGGACGTCATGAACTGTTCACCATCAATACGTT 1437
DB 307 GluLysProTrpAspProPheAspArgSerHisLeuLysLeuPheThrIleuLysVal 326
QY 1438 GATGTCGATTCACCAAGCAATGTCTTAAAGAGACAGCTTTACAGCTAGGAAACGTAAC 1497
DB 327 AspValAsnThrAsnGlnLeuLeuLysSerGlnGlnIleuLeuThrAlaSerGlnArgAsn 346
QY 1498 TTAGCTTCAAGATTTATACGATCTCGTATAGGCTTAACTACTCTCAACAATCTC 1557
DB 347 LeuAspPheArgAspLeuTrpAspProAlaArgAspLysAlaLysLeuLeuTrpAsnAsnLeu 366
QY 1558 GATGCTTTTGGTATTTATGACATATACCTTACTGAGAAAGTGAAGATAATCACAGTAC 1617
DB 367 AspAlaPheGlyIleLeuAspTrpTrpLeuThrGlyLysValGlnAspAsnHisAspAsp 386
QY 1618 ACCAACCGTATCAATACCGTTTATAGGCGAAGACCGCAAGAGGAATGCTAGCTAT 1677
DB 387 ThrAsnArgIleIleThrValIleuMetGlyLysArgProGlnGlnIleuValAlaSerTrp 406
QY 1678 CATTTAGCCTGATTAAGATCGTTATACCGAAGAAAGAGAAAGTTCACGCTACTG 1737
DB 407 HisLeuAlaIleTrpAspLysAspArgTrpThrGlnGlnGlnIleuArgIleuValIleu 426
QY 1738 CGTTATACAGGACACCTTATACCTGATTAACCTTAACGACAA 1779
DB 427 ArgTrpThrGlyThrProIleProAspAsnProAspAsnLys 440

RESULT 13
US-07-854-596B-47
; Sequence 47, Application US/07854596B
; Patent No. 5434073
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith M
; APPLICANT: Hunter, Michael G
; APPLICANT: Czaplowski, Lloyd G
; TITLE OF INVENTION: Proteins and nucleic acids
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. John J. McDonnell
; STREET: Ten South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/854,596B
; FILING DATE: 03-JUN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,337
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 483 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-854-596B-47

```

```

Alignment Scores:
Pred. No.: 1,836-207 Length: 483
Score: 2132.00 Matches: 409
Percent Similarity: 99.03% Conservative: 1
Best Local Similarity: 98.79% Mismatches: 4
Query Match: 68.03% Indels: 0
DB: 1 Gaps: 0

```

US-09-940-235-11 (1-1782) x US-07-854-596B-47 (1-483)

```

QY 538 ATTGCTGAGCTGAGTGGCTGCTAGACCGTCCATCTCTCAACAACACCAATTGGTTGT 597
DB 1 TLeaAGlyProGluTrpLeuLeuAspArgProSerValAsnSerGlnLeuValVal 20
QY 598 AGCGTGTGCTGCTAGTCTGTTGAGGGAACGAATCAAGACATTAGTCTTAATTTTGAATC 657
DB 21 SerValAlaGlyThrValGlnGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 40
QY 658 GATCTAACATCAGACCTGCTCATGAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 717
DB 41 AspLeuThrSerArgProAlaHisGlyGlyLeuThrGlnGlnGlnGlnGlnGlnGlnGlnGln 60
QY 718 AAAACCTTTGCTACTGATAGTGGCGGATGTCATATACTTTGAGAAAGCTGACTTACTA 777
DB 61 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGlnLysAlaAspLeuLeu 80
QY 778 AAGGCTATTCAAGAAATTTGATCGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 837
DB 81 LysAlaIleGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 100
QY 838 ATGTGATTTTGCAGCGATGCAACCATTAAGTCAAGCGCAAGTCTCTTTGCTGAC 897
DB 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValIlyrPheAlaAsp 120
QY 898 AAAAGATGTTCCGTTAACTTGGCCGACCCAACTGCTCCAGAAATTTTGTAAAGCGACAT 957
DB 121 LysAspGlySerValThrLeuProThrGlnProValGlnGlnGlnGlnGlnGlnGlnGlnGln 140
QY 958 GTGCGCGTTAGACCAATATAAGAAAAACAATACAAACCAAGCGAAATCTGTGATGTG 1017
DB 141 ValArgValArgProIlyrLysGlyLysProIleGlnAsnGlnAlaLysSerValAspVal 160
QY 1018 GAATACTACTGACAGTTTACCTCTTAAACCTGATGACGATTTGAGACAGGCTGCAAA 1077
DB 161 GlnIlyrThrValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 180
QY 1078 GATCTAACCTATTGAAAAACATAGTATCGGTGACACCATCTCAAGATTACTA 1137
DB 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeuLeu 200

```

```

QY 1138 GCTCAGACAAAGACATTTTAAACAAAAACCAACCGAGCTATAGATTATGACGTGAC 1197
DB 201 AlaGlnIleGlnSerIleLeuAsnLysThrHisProGlyIlyrThrIleIlyrGlnArgAsp 220
QY 1198 TCTCTCAATCGTCACTGACATGACATGACATTTTCCGTAGATTATTCACATGATCAAG 1257
DB 221 SerSerIleValIThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGln 240
QY 1258 TTTACTTACCGGTAAAAAATCGGACACAGCTTTATAGATCAATAAAAAATCTGCTCG 1317
DB 241 PheThrIlyrHisValLysAsnArgGlnGlnAlaIlyrGlnIleAsnLysLysSerIlyLeu 260
QY 1318 AATGAAAGAAATTAACACACATGACCTGATCTCTGAGAAATATTAAGCTTTAAAAAGGG 1377
DB 261 AsnGlnGlnIleAsnAsnThrAspLeuIleSerGlnLysIlyrIlyrValLeuLysGly 280
QY 1378 GAAAAAGCGTATGATCCTTTGATGCGACGTGACCTTGAACCTTTACCATCAATACGTT 1437
DB 281 GlnLysProIlyrAspProPheAspArgSerHisLeuLysLeuPheThrIleIlyrVal 300
QY 1438 GATGTCGATCAACCAAGATTGCTAAAGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1497
DB 301 AspValAsnThrAsnGlnLeuLeuLysSerGlnGlnLeuThrAlaSerGlnArgAsn 320
QY 1498 TTAGACTTCAGAGATTATTAAGATCTCGTATTAAGCTTAAGCTACTCTACACATCTC 1557
DB 321 LeuAspPheArgAspLeuIlyrAspProArgAspLysAlaLysLeuLysIlyrAsnLeu 340
QY 1558 GATGCTTTGGTATTATGACTATACCTTAACTGAAAAAGTAGAGATATCAAGATGAC 1617
DB 341 AspAlaPheGlyIleMetAspIlyrThrLeuThrGlyLysValGlnAspAsnHisAspAsp 360
QY 1618 ACCAACCGTATCATATACCGTTTATATGAGGCAAGGACCCGAAAGAGAAATGCTAT 1677
DB 361 ThrAsnArgIleIleThrValIlyrMetGlyLysArgProGlnGlnGlnGlnGlnGlnGlnGlnGln 380
QY 1678 CATTTAGCGTATGATTAAGATCGTTATACCGAAGAAAGAGAAAGTTTACAGCTACCTG 1737
DB 381 HisLeuAlaIlyrAspLysAspArgIlyrThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 400
QY 1738 CGTTATACGAGACACCTATACCTGATTAACCTTAACGACAA 1779
DB 401 ArgIlyrThrGlnThrProIleProAspAsnProAsnAspLys 414

RESULT 14
US-07-854-596B-35
; Sequence 35, Application US/07854596B
; Patent No. 5434073
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith M
; APPLICANT: Hunter, Michael G
; APPLICANT: Czaplowski, Lloyd G
; TITLE OF INVENTION: Proteins and nucleic acids
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. John J. McDonnell
; STREET: Ten South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/854,596B
; FILING DATE: 03-JUN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949

```



```

REFERENCE/DOCKET NUMBER: 92,337
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
TELEX: 910-221-5317
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 859 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-854-596B-35

Alignment Scores:
Pred. No.: 2,44e-207 Length: 859
Score: 2132.00 Matches: 409
Percent Similarity: 99.03% Conservative: 1
Best Local Similarity: 98.79% Mismatches: 4
Query Match: 68.03% Indels: 0
DB: 1 Gaps: 0

US-09-940-235-11 (1-1782) x US-07-854-596B-35 (1-859)
QY 538 ATTGCTGACCTGATGCTGCTAGACCGTCCATCTGTCAACAACAGCCAAATTGTTGTT 597
DB 22 TleaaGlyProGluThrIleuLeuAspArgProSerValAsnAsnSerGlnLeuVal 41
QY 598 AGCGTGTGCTAGCTGTGAGGGAGCAATCAAGACATTAGTCTTAATTTTGAATC 657
DB 42 SerValaIaGlyThrValGluGluThrAsnGlnAspIleSerIleuLysPheGluIle 61
QY 658 GATCTAACATCCAGACCTGCTCATGTGAGAGAAAGACAGACAGCGCTTAAGTCCAAATCA 717
DB 62 AspleuThrSerArgProAlaHisGlyGlyLysThrGluGluGlyLeuSerProLysSer 81
QY 718 AAACCAATTTGCTACGATGATGCGCGAGATGTCAACATTAACCTTGAGAAAGCTGACTACTA 777
DB 82 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu 101
QY 778 AAGGCTATTCAAGAACATTTGATCGTAACTGCAACAGTACAGACACTACTTTGAGTTC 837
DB 102 LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 121
QY 838 ATTGATTTTGCAGGAGTCAACCATTTACTGATTCGAACCGCAAGGTCTACTTGTCTGAC 897
DB 122 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 141
QY 898 AAAGATGCTGGTGAACCTTGGCGAACCCCAACCTGTCCAGAAATTTTGTACCGGACAT 957
DB 142 LysAspGlySerValThrIleuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 161
QY 958 GTGCGCGTTAGACCATATTAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTC 1017
DB 162 ValArgValArgProTyrIleLysGlyLysProIleGlnAsnGlnAlaLysSerValAspVal 181
QY 1018 GAATTACTGTACAGTTACTCCCTTAAACCTGATGACGATTTCCAGCCAGGCTCTCAA 1077
DB 182 GluTyrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuLys 201
QY 1078 GATCTAGCTATTGAAAAACATAGCTATCGGTGACACCATCAACATCTCAAGAAATTAACA 1137
DB 202 AspThrLysLeuLeuLysThrIleuValIleGlyAspThrIleThrSerIleGlnLeuLeu 221
QY 1138 GCTCAAGACAAGACATTTTAAACAAAACCAACCCAGGCTATACGATTATGAAAGCTGAC 1197
DB 222 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleIleTyrGluAspArg 241
QY 1198 TCTTCATCTGCTACATGACATGACATGATTTCCGATGATTTTCAATGATCAAGAG 1257
DB 242 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 261
QY 1258 TTTACTTACCGGTGTTAAAAATGGGAAACAGCTTATAGATCAATAAAAATCTGGTCTG 1317

```

```

DB 262 PheThrTyrHisValLysAsnArgGluGlnAlaIleTyrGluIleAsnLysLysSerGlyLeu 281
QY 1318 AATGAAGAAATAATAACAACATGACCTGATCTCTGAGAAATATATACGCTTAAAAAGG 1377
DB 282 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 301
QY 1378 GAAAGCCCGATATGATCCCTTTGATCGCAGTCACTTGAAACGTTCCATCAATAACGTT 1437
DB 302 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 321
QY 1438 GATGTGATACCAACGAATTGCTAAAAAGTACAGCAGCTCTTAACAGCTACGAAAGTAA 1497
DB 322 AspValAsnThrAsnGlnLeuLeuLysSerGluGlnLeuThrAlaSerGluAsn 341
QY 1498 TTTAGCTTCAGACATTTATTCGATCTCCGATTAAGGCTTAATACTCTACAAATCTC 1557
DB 342 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLysTyrAspAsnLeu 361
QY 1558 GATGCTTTTGGTATTTATGACTATACCTTAACCTGGAAGAGTAAAGATTAACGATGAC 1617
DB 362 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 381
QY 1618 ACCAACCGTATCATTAACCGTTTATATGGCGAACCGAACCCGAAAGAGAAATGCTAGCTAT 1677
DB 382 ThrAsnArgIleIleThrValTyrMetGlyLysArgProIleGluGluAsnAlaSerTyr 401
QY 1678 CATTAGCTTATGATTAAGATCGTTATATCCGAAGAAAGACGAAAGTTTACAGCTACTG 1737
DB 402 HisLeuAlaTyrAspLysAspArgTyrThrGluGluGluValTyrSerTyrLeu 421
QY 1738 CGTTATACAGGAGCACCTTAACCTGATTAACCTTAACGTAACGACGACAA 1779
DB 422 ArgTyrThrGlyThrProIleProAspAsnProAsnAspLys 435

RESULT 15
US-08-488-940-17
; Sequence 17, Application US/08488940
; Patent No. 5854049
; GENERAL INFORMATION:
; APPLICANT: Reed, Guy L.
; TITLE OF INVENTION: PLASMIN-RESISTANT STREPTOKINASE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,940
; FILING DATE: 09-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 05433/009001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1194 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein

```

US-08-488-940-17

Alignment Scores:

Pred. No.:	4, 24e-206	Length:	1194
Score:	2120.50	Matches:	415
Percent Similarity:	96.52%	Conservative:	1
Best Local Similarity:	96.29%	Mismatches:	10
Query Match:	67.66%	Indels:	5
DB:	1	Gaps:	3

US-09-940-235-11 (1-1782) x US-08-488-940-17 (1-1194)

```

QY 496 CAGACCAATCCAGC-----GGATCTGGCCCTTCACCGATGTTGTAATGCTGA 546
DB 366 GlnThrAspSerSerSerValProGlyArgGlySerIle---GlnGlyArgGlyLeuIleGly 384
QY 547 CCGATGTCGTGTATACCGTCATCTGTCACAAAGCCCAATGTTGTTAGGCTGCT 606
DB 385 ProGlnTyrLeuLeuSerPalaProSerValAsnAsnSerGlnLeuValIleSerValIle 404
QY 607 GGTACTGTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAATCGATCTACA 666
DB 405 GlyThrValGlnGlyThrAsnGlnAspIleSerLeuAlaPhePheGlnIleAspLeuThr 424
QY 667 TCACGACCTGCTCATGAGGAAAGACAGCAAGGCTTAAAGTCCAAATCAAAACCATTT 726
DB 425 SerAlaProAlaHisGlyGlyAlaThrGlnGlnGlyLeuSerProAlaSerIleProPhe 444
QY 727 GGTACTGATAGTGGCCGATGTCACTAACTTGAGAAAGCTGACTTAAAGGCTAAT 786
DB 445 AlaThrAspSerGlyAlaMetSerHisIleGlnGlnIleAspLeuLeuValIle 464
QY 787 CAAGAAACAATTGATGCTTAACGTCACAGTAAGACGACTTGAAGTCATTAATTTT 846
DB 465 GlnGlnGlnLeuLeuIleAsnValHisSerAsnAspAspTyrPheGlnValIleAspPhe 484
QY 847 GCAAGCGATGCAACCATTAAGTCAAGAAAGGCAAGGCTTACTTGTGACAAAGATGT 906
DB 485 AlaSerAspAlaThrIleThrAspArgAsnGlyIleValIleThrPheAlaAspIleAspGly 504
QY 907 TTGGTAACTTGGCCGACCCGACCTGTCCAAAGATTTTGGTAAAGCGGACATGTGGCGGT 966
DB 505 SerValThrLeuProThrGlnProValGlnGlnPheLeuLeuSerGlyHisValArgVal 524
QY 967 AGACCATTAATAAGAAAACCAATACAAACCAAGCCGAAATCTGTGATGTGAATATATCT 1026
DB 525 Arg---TyrLysGlnLysProIleGlnAsnGlnAlaLysSerValAspValGlnTyrThr 543
QY 1027 GTACAGTTTACTCCCTTAACCTGATGACGATTTTCAGACGAGTCTCAAGATATCTAAG 1086
DB 544 ValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuLysAspThrLys 563
QY 1087 CTATTGAAAACCTAGCTATCGGTGACACCATCACATCTTCAGAAATTAAGTCTCAAGCA 1146
DB 564 LeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeuLeuAlaGlnAla 583
QY 1147 CAAGCATTTTAAACAAAACCAAGGCTATAGATTTATGAAAGTGAAGTCTCTCAATC 1206
DB 584 GlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGlnLysArgAspSerSerIle 603
QY 1207 GTCACTCATGACATGACATTTTCCGTAGATTTTACCAATGATGACAGGTTTACTTAC 1266
DB 604 ValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlnPheThrTyr 623
QY 1267 CGTGTAAATAATCGGGAACAAAGCTTATAGATCAATAAAAAATCTGGTCTGAATGAAGA 1326
DB 624 ArgValLysAsnArgGlnGlnAlaTyrArgIleAsnLysLysSerGlyLeuAsnGlnGln 643
QY 1327 ATAAACAACACTGACCTGATCTCTGAGAAATATTAAGTCTTAAAAAGGGGAAAAAGCCG 1386
DB 644 IleAsnAsnThrAspLeuIleSerGlnLysTyrTyrValLeuLysLysGlyGlnLysPro 663
QY 1387 TATGATCCCTTTCATCGACGTCACTTGAAACTGTTCACATCAATAACGTTGATGTGAT 1446

```

```

DB 664 TyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrValAspValAsp 683
QY 1447 ACCAAGCAATTCGTTAAAGTAGAGACGCTTAAACGCTTACGGAACGTAACCTTACCTTC 1506
DB 684 ThrAsnGlnLeuLeuLysSerGlnGlnLeuLeuThrAlaSerGlnArgAsnLeuAspPhe 703
QY 1507 AGAGATTATAGATCCGTCGATPAGGCTTAAAGCTTAACTCTTACCAACAATCGATGCTTTT 1566
DB 704 ArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeuAspAlaPhe 723
QY 1567 GGTATTATGACCTATACCTTAACTGGAAGTAGAGATTAATCACAGATGACACCAACCGT 1626
DB 724 GlyIleMetAspTyrThrLeuThrGlyLysValGlnAspAsnHisAspAspThrAsnArg 743
QY 1627 ATCATTAACGCTTTATATGAGGCAAGCAAGCCGGAAGAGAAATGCTAGCTATATTTAGCC 1686
DB 744 IleIleThrValTyrMetGlyLysArgProGlnGlnGlnAsnAlaSerTyrHisLeuAla 763
QY 1687 TATGATTAAGATCGTTATACCGAAGAAAGAGAGATTTACAGCTACCTGGTATACA 1746
DB 764 TyrAspLysAspArgTyrThrGlnGlnGlnIleArgGlnValIleTyrSerTyrLeuArgTyrThr 783
QY 1747 GGACACCTATACCTGATTAACCTTAACGACAAA 1779
DB 784 GlyThrProIleProAsnProAsnProAsnLys 794

```

Search completed: January 28, 2006, 02:41:40
 Job time : 48.4212 secs

GenCore version 5.1.6
 Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 28, 2006, 01:51:16 ; Search time 116.535 Seconds
 (without alignments)
 12778.538 Million cell updates/sec

Title: US-09-940-235-11

Perfect score: 3134
 Sequence: 1 tcgccttcacgtcgtcgtcg.....ataacctaacgacaataa 1782

Scoring table:
 BLOSUM62
 Xgapop 10.0 , Xgapext 0.5
 Ygapop 10.0 , Ygapext 0.5
 Fgapop 6.0 , Fgapext 7.0
 Delop 6.0 , Delext 7.0

Searched: 1867569 segs, 417829326 residues

Total number of hits satisfying chosen parameters: 3735138

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Command line parameters:

MODEL=frame_n2p.model -DEV=slp
 O=/cgn2_1/USPTO_bpool_p/US09940235/runat_27012006_144219_27635/app_query.fasta_1.7708
 -DB=Published Applications AA Main -OPMT=fastan -SUFFIX=rapbm -MINMATCH=0.1
 -LOOPEXT=0 -UNITS=bites -START=1 -MATRIX=blosum62
 -TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
 -THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTPMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0
 -MAXLEN=2000000000 -USRR=US09940235 @CGN_1_1.805 @runat_27012006_144219_27635
 -NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NSG_SCORES=0 -WAIT -DSPBLOCK=100
 -LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
 -FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA Main:*

- 1: /cgn2_6/prodata/1/pubppaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/prodata/1/pubppaa/US08_PUBCOMB.pep.*
- 3: /cgn2_6/prodata/1/pubppaa/US09_PUBCOMB.pep.*
- 4: /cgn2_6/prodata/1/pubppaa/US10_PUBCOMB.pep.*
- 5: /cgn2_6/prodata/1/pubppaa/US10_PUBCOMB.pep.*
- 6: /cgn2_6/prodata/1/pubppaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2158	68.9	414	3	US-09-940-235-2
2	2153	68.7	413	4	US-10-360-101-264
3	2132	68.0	414	4	US-10-300-215-252
4	2132	68.0	415	5	US-10-988-943-1
5	2121	67.7	414	4	US-10-300-215-253
6	2090.5	66.7	413	3	US-09-919-703-12
7	1937	61.8	440	5	US-10-474-792-658
8	504.5	16.1	2320	4	US-10-236-392-2
9	504.5	16.1	2328	4	US-10-171-311-64
10	504.5	16.1	2328	4	US-10-236-031B-70
11	504.5	16.1	2328	4	US-10-374-979-98
12	504.5	16.1	2328	4	US-10-182-936A-98

13	504.5	16.1	2328	5	US-10-477-238A-677
14	504.5	16.1	2328	5	US-10-680-287A-677
15	504.5	16.1	2328	5	US-10-477-173-677
16	503	16.0	463	4	US-10-144-194A-52
17	503	16.0	463	5	US-10-491-566-52
18	503	16.0	642	4	US-10-741-601-354
19	503	16.0	642	5	US-10-741-600-1066
20	503	16.0	657	4	US-10-741-601-359
21	503	16.0	657	5	US-10-741-600-1072
22	503	16.0	984	4	US-10-741-601-356
23	503	16.0	984	5	US-10-741-600-1069
24	503	16.0	2220	4	US-10-236-392-4
25	503	16.0	2296	4	US-10-741-601-363
26	503	16.0	2296	5	US-10-741-600-1075
27	503	16.0	2320	4	US-10-279-733-8
28	503	16.0	2355	4	US-10-144-194A-104
29	503	16.0	2355	4	US-10-360-101-235
30	503	16.0	2355	4	US-10-447-161-3
31	503	16.0	2355	4	US-10-734-564-94
32	503	16.0	2355	4	US-10-741-601-357
33	503	16.0	2355	4	US-10-741-601-366
34	503	16.0	2355	5	US-10-491-566-104
35	503	16.0	2355	5	US-10-741-600-1067
36	503	16.0	2355	5	US-10-741-600-1078
37	503	16.0	2355	5	US-10-852-335A-147
38	503	16.0	2355	5	US-10-287-436A-436
39	503	16.0	2355	5	US-10-287-436A-1137
40	503	16.0	2355	6	US-11-040-130-28
41	503	16.0	2386	3	US-09-961-403-1
42	503	16.0	2386	4	US-10-741-601-360
43	503	16.0	2386	4	US-10-788-792-206
44	503	16.0	2386	5	US-10-618-281-32
45	503	16.0	2386	5	US-10-741-600-1071

Sequence 677, App
Sequence 677, App
Sequence 677, App
Sequence 52, App1
Sequence 52, App1
Sequence 354, App
Sequence 1066, App
Sequence 359, App
Sequence 1072, App
Sequence 356, App
Sequence 1069, App
Sequence 363, App1
Sequence 363, App1
Sequence 8, App1
Sequence 104, App
Sequence 235, App
Sequence 3, App1
Sequence 94, App1
Sequence 357, App
Sequence 366, App
Sequence 1067, App
Sequence 1067, App
Sequence 147, App
Sequence 436, App
Sequence 1137, App
Sequence 28, App1
Sequence 1, App1
Sequence 360, App
Sequence 206, App
Sequence 32, App1
Sequence 1071, App

ALIGNMENTS

RESULT 1
 US-09-940-235-2
 ; Sequence 2, Application US/09940235
 ; Publication No. US20030059921A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kumar, Rajesh
 ; APPLICANT: Sahni, Girish
 ; APPLICANT: Roy, Chait
 ; APPLICANT: Rajagopal, Kammarra
 ; APPLICANT: Nihalani, Deepak
 ; APPLICANT: Sundaram, Vasudha
 ; APPLICANT: Yadav, Mahavir
 ; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
 ; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
 ; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
 ; FILE REFERENCE: 07064-009002
 ; CURRENT APPLICATION NUMBER: US/09/940,235
 ; CURRENT FILING DATE: 2002-04-09
 ; PRIOR APPLICATION NUMBER: 09/471,349
 ; PRIOR FILING DATE: 1999-12-23
 ; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
 ; PRIOR FILING DATE: 1998-12-24
 ; NUMBER OF SEQ ID NOS: 28
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 414
 ; TYPE: PRT
 ; ORGANISM: Streptococcus equisimilis
 US-09-940-235-2
 Alignment Scores:
 Pred. No.: 3.4e-192
 Score: 2158.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Length: 414
 Matches: 414
 Conservative: 0
 Mismatches: 0

Query Match: 68.86% Indels: 0
 DB: 3 Gaps: 0
 US-09-940-235-11 (1-1782) x US-09-940-235-2 (1-414)

```

QY 538 ATTGCTGACCTGAGTGCGCTGTAGACCGTCATCTGTCAACAACCAAGTGGTTGT 597
Db 1 IleaIagIyProGIuThrIeuAaPArGProSeVaIaSnnsSergInIeuVal 20
QY 598 AGCGTTCGTGACTGTGAGGGAGACGATTCAGACATTAGTTAAATTTTGAATC 657
Db 21 SerVaIaIagIyThrValGIuGIuThraNGInAspIleSeRleuysPhePheGluIle 40
QY 658 GATCTAACATCAGACCTGCTCATGAGGAAAGACAGAGAGGCTTAAGTCCAAATCA 717
Db 41 AspleuThSeRarProIaHIsGIyGIyIysThGIuGIuInIySeuSeRProIySe 60
QY 718 AAACCATTTGCTACTGATAGTGCGCGGATGTCAATMACTTGAGAAAGCTGACTTA 777
Db 61 LysPProPheIaThraSperGIyAlaMeSeRHisIyIeuGIuIySaIaAspleu 80
QY 778 AAGGTATTTCAAGAACATTTGCTTACGTTACAGTACAGACATCTTGAAGGTC 837
Db 81 LysaIaIleGIuGIuInIeuIleaIaSnValHisSeRAsnaSparGIyRPhGIuVal 100
QY 838 ATTGATTTTGAAGGATGCAACCATTAAGTCAAGGAGGCTTAAGTTCGTCGAC 897
Db 101 IleaSperPheIaSeRAspaIaThrIleThraSparGAnsGIyIysValIyRPhIaAsp 120
QY 898 AAAGATGGTTCGTGAACCTTGCGACCAACCTGTCAGAAATTTTGGTAAAGCGACAT 957
Db 121 LysaSperGIySeRValThrIeuProthGIuInProValGIuGIuPheIeuSeuSeRGIyHis 140
QY 958 GTGCGCGTTAGACCATTAAGAAAAACCAATCAAAACCAAGCGAATCTGTTGATGTG 1017
Db 141 ValaRgValaIarGProIyRGIySGLIySProIleGIuInaGIuIaIySeRValaIaspaI 160
QY 1018 GAATTAACCTGTAAGATTTTACCTTAAACCTGATGAGATTTTGACAGCTTCACAA 1077
Db 161 GIuIyThrIValaGIuPheThraProIeuaInProIaSparSparPheThraGIyIeuIyS 180
QY 1078 GATACTAAGCTATTGAAAAACATAGCTATCGGTGACACATCATCTCAAGAAATCTA 1137
Db 181 AspThrIySleuIeuIyThrIeuIaIleGIyAspThrIleThSeRGIuGIuIeu 200
QY 1138 GCTCAAGCAAAAGATTTTAAACAAAAACCAACCGGCTTAACGATTTAAGAACGTGAC 1197
Db 201 AlaGIuIaGIuInSeRilleuSnIySaSnHisPProGIyIyThrIleThRGIuIaIyGaR 220
QY 1198 TCCCTCAATCTGCATGACGACATGATTTCCGTAAGATTTTACCATGATGACAAAG 1257
Db 221 SeRSeRilleValThrHisAspaSnAspIlePheIaGIyThrIleuProMeSeRGIuGIu 240
QY 1258 TTTACTTACCGTGTAAATAAATCGGGAACAAGCTTATAGGATCAATAAAATCGGTG 1317
Db 241 PheThrIyRgValaIySaSnIarGIuGIuIaIyRgIleuSnIySSeRGIyIeu 260
QY 1318 AATGAAGAATTAACAACATGACCTGATCTGTGAGAAATATTACGTCTTAAAAAGG 1377
Db 261 AsnGIuGIuIleAsnaSnThraSperIleSeRGIuIyIyThrIleuIySGLIy 280
QY 1378 GAAAAAGCGGTATGATCCCTTTGATGCGACATCACTTGAATCTGTCACCAATACGTT 1437
Db 281 GIuIySProIyRAsPProPheAspaRSeRHisIleuIySleuPheThrIleuIyRVal 300
QY 1438 GATGTCGATTAACAAGAAATTTGTAATAAGTGACGACTTAAACACTTGACGAAAGCTAAC 1497
Db 301 AspaIaAspThraSngIuIeuIeuIySeRGIuGIuInIeuIyThraSseRGIuIaIyGaR 320
QY 1498 TTAGACTTCAGAGATTTATACGATCTCGTGAATAAGGCTTAACTACTTACAAATCTC 1557
Db 321 LeuAsPheIaRgAspleuIyThraSProIaRgAspIySaIaIySleuIeuIyThraSnaIeu 340

```

```

QY 1558 GATGCTTTTGGTATTATGACTATACCTTAAGTGAAGAGATGAATACAGATGAC 1617
Db 341 AspaIaPheIyIleIeMeSaRyThrIleThGIySaIaGIuAsnaSnIaAspaR 360
QY 1618 ACCAACCGTATCATTAACCGTTTATATGGGACAGGACCCGAAAGAGATGCTAGTAT 1677
Db 361 ThraSnaRGIleIleThrIValIyMeGIyIySaRGIySeRProGIuGIuInaIaSeRy 380
QY 1678 CATTGACCTATGATTAAGATCGTTATACCGAAGAAAGAGAGATTTACAGTACTCG 1737
Db 381 HisIeuIaIyRAspIySaRGIyThrGIuGIuInGIuIyRgIuIaIySeRyIeu 400
QY 1738 CGTTATACAGGACACCTATACCTGATTAACCTTAACGACAA 1779
Db 401 ArgIyThrGIyThrProIleProIaSparSparSnaSnaIyS 414

RESULT 2
US-10-360-101-264
; Sequence 264, Application US/10360101
; Publication No. US20040009550A1
; GENERAL INFORMATION:
; APPLICANT: MoII, Gert N.
; APPLICANT: Leenhouts, Cornelis J.
; TITLE OF INVENTION: Export and modification of (poly)peptide in the lantibiotic way
; FILE REFERENCE: 2183-5673
; CURRENT APPLICATION NUMBER: US/10/360,101
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: EP 02077060.8
; NUMBER OF SEQ ID NOS: 309
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 264
; LENGTH: 413
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: sequence of streptokinase
US-10-360-101-264

Alignment Scores:
Pred. No.: 9.97e-192 Length: 413
Score: 2153.00 Matches: 413
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 68.70% Indels: 0
DB: 4 Gaps: 0

US-09-940-235-11 (1-1782) x US-10-360-101-264 (1-413)

QY 538 ATTGCTGACCTGAGTGCGCTGTAGACCGTCATCTGTCAACAACCAAGTGGTTGT 597
Db 1 IleaIagIyProGIuThrIeuAaPArGProSeVaIaSnnsSergInIeuVal 20
QY 598 AGCGTTCGTGACTGTGAGGGAGACGATTCAGACATTAGTTAAATTTTGAATC 657
Db 21 SerVaIaIagIyThrValGIuGIuThraNGInAspIleSeRleuysPhePheGluIle 40
QY 658 GATCTAACATCAGACCTGCTCATGAGGAAAGACAGAGAGGCTTAAGTCCAAATCA 717
Db 41 AspleuThSeRarProIaHIsGIyGIyIysThGIuGIuInIySeuSeRProIySe 60
QY 718 AAACCATTTGCTACTGATAGTGCGCGGATGTCAATMACTTGAGAAAGCTGACTTA 777
Db 61 LysPProPheIaThraSperGIyAlaMeSeRHisIyIeuGIuIySaIaAspleu 80
QY 778 AAGGTATTTCAAGAACATTTGCTTACGTTACAGTACAGACATCTTGAAGGTC 837
Db 81 LysaIaIleGIuGIuInIeuIleaIaSnValHisSeRAsnaSparGIyRPhGIuVal 100
QY 838 ATTGATTTTGAAGGATGCAACCATTAAGTCAAGGAGGCTTAAGTTCGTCGAC 897
Db 101 IleaSperPheIaSeRAspaIaThrIleThraSparGAnsGIyIysValIyRPhIaAsp 120

```

```

QY 898 AAAAGATGCTTCGGTAACTTCCGACCCAACTGTCCAGAAATTTTGGTAAAGCGGACAT 957
    |||||
Db 121 LysAspGlySerValThrLeuProThnGlnProValGlnGlnPheLeuSerGlyHis 140
QY 958 GTGCGGGTTAGACCATTAAGAAAAACAATACAAAACCAAGGAAATCTGTGATGTC 1017
    |||||
Db 141 ValArgValArgProTyrLysGlnLysProIleGlnAsnGlnAlaLysSerValAspVal 160
QY 1018 GAATATACTGTACAGTTACTCCCTTAAACCCGTATGACGATTTCAACAGCTCTCAAA 1077
    |||||
Db 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
QY 1078 GATCTAAGCATTTGAAAAACACTAGCTATCGGTGACCAATCATCATCTCAAGAAATTA 1137
    |||||
Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeu 200
QY 1138 GCTCAAGCAAAAGCAATTTTAAACAAAACCAACCCAGGCTATACAGATTTATGAACGTGAC 1197
    |||||
Db 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGlnArgAsp 220
QY 1198 TCCCTAATGCTCACTCATGACAAATGACATTTTCCGTACAGATTTTACCAAGATCAAGAG 1257
    |||||
Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGln 240
QY 1258 TTTACTTACCGTGTAAATAATCGGAAACAAGCTTATAGATCAATAAAAAATCGTCTG 1317
    |||||
Db 241 PheThrTyrArgValLysAsnArgGlnGlnAlaTyrArgIleAsnLysLysSerGlyLeu 260
QY 1318 AATGAAGAATAAACAACACTGACCTGATCTTGAGAAATATTAAGCTTTAAAAAGG 1377
    |||||
Db 261 AsnGlnGlnLysIleAsnAsnThrAspLeuIleSerGlnLysTyrTyrValIleLysLysGly 280
QY 1378 GAAAGACCGTATGATCCCTTGTATGCGAGTCACTTGAAAGCTTGCACCAATTAAGTT 1437
    |||||
Db 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
QY 1438 GATGTCGATACCAACAATTTGCTAAAGAGTACAGCTTTAAACAGCTACGAAACGTAAC 1497
    |||||
Db 301 AspValAspThrAsnGlnLeuLeuLysSerGlnGlnLeuThrAlaSerGlnLysArgAsn 320
QY 1498 TTAGACTTCAGAGATTTATACGATCTCTGTATAGGCTTAAGCTTAACAAATCTC 1557
    |||||
Db 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLysTyrAsnAsnLeu 340
QY 1558 GATGCTTTGGTATTAATGACATTAACCTTAACGAAAGATAGAGATTAACAGATGAC 1617
    |||||
Db 341 AspAlaAspGlyLysLeuAspTyrThrLeuThrGlyLysValGlnAspAsnHisAspAsp 360
QY 1618 ACCAACCGTATCAATACCGTTATATGAGGACGACCCGAGAGAGAGAAATGCTAGCTAT 1677
    |||||
Db 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGlnGlnLysAsnAlaSerTyr 380
QY 1678 CATTTAGCCTATGATTAAGATCGTTATACCGAAGAAAGAAAGAAAGTTTACAGTACTG 1737
    |||||
Db 381 HisLeuAlaTyrAspLysAspArgTyrThrGlnGlnGlnArgLysValTyrSerTyrLeu 400
QY 1738 CGTTATACGGAACACTTATACCTGATTAACCTTAACGAC 1776
    |||||
Db 401 ArgTyrThrGlyThrProIleProAspAsnProAspAsn 413

```

RESULT 3

```

US-10-300-215-252
; Sequence 252, Application US/10300215
; Publication No. US20030153043A1
; GENERAL INFORMATION:
; APPLICANT: CARR, Francis Joseph
; APPLICANT: ADAIR, Fiona Suzanne
; APPLICANT: HAMILTON, Anita Anne
; APPLICANT: CARTER, Graham
; TITLE OF INVENTION: METHOD FOR THE PRODUCTION OF
; TITLE OF INVENTION: NON-IMMUNOGENIC PROTEINS
; FILE REFERENCE: MER-104-Con.1
; CURRENT APPLICATION NUMBER: US/10/300,215

```

```

; CURRENT FILING DATE: 2002-11-20
; PRIOR APPLICATION NUMBER: US 09/438,136
; PRIOR FILING DATE: 1999-11-10
; PRIOR APPLICATION NUMBER: WO PCT/GB98/01473
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: GB 9710480.6
; PRIOR FILING DATE: 1997-05-21
; PRIOR APPLICATION NUMBER: GB 9716197.0
; PRIOR FILING DATE: 1997-07-31
; PRIOR APPLICATION NUMBER: GB 9725270.4
; PRIOR FILING DATE: 1997-11-28
; PRIOR APPLICATION NUMBER: GB 9807751.4
; PRIOR FILING DATE: 1998-04-14
; PRIOR APPLICATION NUMBER: US 60/067,235
; PRIOR FILING DATE: 1997-12-02
; NUMBER OF SEQ ID NOS: 254
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 252
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Streptococcus equisimilis
US-10-300-215-252
Alignment Scores:
Pred. No.: 9,12e-190 Length: 414
Score: 2132.00 Matches: 409
Percent Similarity: 99.03% Conservative: 1
Best Local Similarity: 98.79% Mismatches: 4
Query Match: 68.03% Indels: 0
Gaps: 0
US-09-940-235-11 (1-1782) x US-10-300-215-252 (1-414)
QY 538 ATTGCTGACCTGATGCTGCTGACACCGTCATCTGCAACAACAGCAATGATGTT 597
    |||||
Db 1 IleAlaGlyProGlnTyrPheLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
QY 598 AGCGTTCGTGCTGCTGTTGGGGGACCAATCAAGACATTAAGCTTTTGAATC 657
    |||||
Db 21 SerValAlaGlyThrValGlnGlyThrAsnGlnAspIleSerLeuLysPheGlnLys 40
QY 658 GATTAACATCAACGACCTGCTCATGAGAGAAAGACAGCAAGGCTTAAGTCAAAATCA 717
    |||||
Db 41 AspLeuThrSerArgProAlaHisGlyGlyLysThrGlnGlnLysSerProLysSer 60
QY 718 AAACCATTTGCTAGTATGAGTGGCGGATGTCACATTAACCTTGAAGAGCTGACTTA 777
    |||||
Db 61 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGlnLysAlaAspLeu 80
QY 778 AAGCTATTCAGAAACAATGATGCTTAACGTCACAGTAACGACGACTTTGAGGTC 837
    |||||
Db 81 LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGlnVal 100
QY 838 ATTGATTTTGAACGATGACCAACATTAAGTATGATGGAACGGCAAGTCTACTTGTGAC 897
    |||||
Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
QY 898 AAAGATGCTTCGTTAAGCTTCCGACCCAACTGTCCAAAGAAATTTTGTAAAGCGGACAT 957
    |||||
Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnGlnPheLeuSerGlyHis 140
QY 958 GTGCGGGTTAGACCATTAAGAAAAACAATACAAAACCAAGGAAATCTGTGATGTC 1017
    |||||
Db 141 ValArgValArgProTyrLysGlnLysProIleGlnAsnGlnAlaLysSerValAspVal 160
QY 1018 GAATATACTGTACAGTTACTCCCTTAAACCCGTATGACGATTTTCAACAGCTCTCAAA 1077
    |||||
Db 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
QY 1078 GATCTAAGCATTTGAAAAACACTAGCTATCGGTGACCAATCATCATCTCAAGAAATTA 1137
    |||||
Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeu 200

```

QY 1138 GCTCAAGCAAAAGATTAAAAACAACCCAGGCTATACGATTATGAACGTGAC 1197
Db 201 AAGlnAlaGlnserlleuLeuLnlysrThrlsProGlytrThrlleryGluArgAsp 220
QY 1198 TCCTCAATCGTACATCAATGACATTTCCGATGATTTTACCAATGATCAAGAG 1257
Db 221 SerSerlleValThrHrlsAspAsnspIlePheArgThrilleuProMetAspGlnGlu 240
QY 1258 TTACTTACCGGTAAAAATCGGGAACAAGCTTATAGGATCAATAAAAAATCGGTG 1317
Db 241 PheThrlyrHrlsVallyAsnAsnArgGlnGlnAlaTyGlnlleAsnlyslsSerGlyLeu 260
QY 1318 AATGAAGAATTAACAACAACCTGATCTGTGAGAAATATTATCGTCTTAAAAAGG 1377
Db 261 AsnGlnGlnlleAsnAsnThrAspIleSerGlnlyTyTrlyValleuLyslsGly 280
QY 1378 GAAAAGCCGTATGATCCCTTTGATCGCACTGCTTGAACCTGTTCACCAATACGTT 1437
Db 281 GluLyrsProTyrsAspProPheAspArgSerHrlsleuLyslsleuPheThrilleuTyTrVal 300
QY 1438 GATGTCGATACCAAGAAATTCGTAATAAAGTGAGCAGCTTAAACAGTCAAGCTAAC 1497
Db 301 AspValAsnThrAsnGlnleuLeuLyserGlnGlnleuThrAlaSerGlnArgAsn 320
QY 1498 TTAGACTTCAGAGATTTATACGATCCCTGATTAAGGCTAAACTACTTACAACAATCTC 1557
Db 321 LeuAspPheArgAspLeuTyrsAspProArgAspLyAlaLyslsleuLeuTyrsAsnleu 340
QY 1558 GATGCTTTTGGTATTATGACATTAACCTTAACCTGAAAAGTGAAGATTAATCAAGTAC 1617
Db 341 AspAlaPheGlylleMetAspTyThrleuThrGlyValGlnAspAsnHrlsAspAsp 360
QY 1618 ACCAACCGTATACATACCGTTATATATGGGCAAGCCAGCAAGAGAGAAATGTTACTAT 1677
Db 361 ThrAsnAspGlnlelleThrVallyrMetGlylsArgProGlnGlyGlnAsnAlaSerTy 380
QY 1678 CATTTAGCCTATGATTAAGATCGTTATACCGAAGAAAGAGAAATTTACAGCTACG 1737
Db 381 HlsleuAlaTyrsAspLyAspArgTyThrGlnGlnGlnAlaVallyrSerTyrlleu 400
QY 1738 CGTTATACAGGACACCTTATACCTGATTAACCTTAACGACAAA 1779
Db 401 ArgTyThrGlyThrProIleProAspAsnProAsnAspLy 414

RESULT 4

US-10-988-943-1
; Sequence 1, Application US/10988943
; Publication No. US20050176085A1
; GENERAL INFORMATION:
; APPLICANT: Center for Genetic Engineering and Biotechnology
; TITLE OF INVENTION: METHOD OF SELECTIVE PEPTIDE ISOLATION FOR THE
; TITLE OF INVENTION: IDENTIFICATION AND QUANTITATIVE ANALYSIS OF PROTEINS IN
; FILE REFERENCE: Proteomics CU2003-269
; CURRENT APPLICATION NUMBER: US/10/988,943
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 415
; TYPE: PRT
; ORGANISM: Streptococcus equisimilis
US-10-988-943-1

Alignment Scores:

Pred. No.: 9,136-190 Length: 415
Score: 2132.00 Matches: 409
Percent Similarity: 99.03% Conservative: 1
Best Local Similarity: 98.79% Mismatches: 4
Query Match: 68.03% Indels: 0
Gaps: 0

US-09-940-235-11 (1-1782) x US-10-988-943-1 (1-415)

QY 538 ATTGCTGAACCTGAGTGGCTGTAGACCGTTCATCTGTCAACAACAGCAATGGTGT 597
Db 2 IleAlaGlyProGlnutrlPleuLeuAspArgProSerValAsnAsnSerGlnleuVal 21
QY 598 AGCGTGTCTGTACTGTGAGGGGACGATCAAGACATTAAGTCTTAAATTTTGAATC 657
Db 22 SerValAlaGlyThrValGlnGlyThrAsnGlnAspIleSerleuLyrsPhePheGlnle 41
QY 658 GATCTAACATCAAGACCTGTCTATGAGAGAAACAGAGCAAGCTTAAAGTCCAAATCA 717
Db 42 AspLeuThrSerArgProAlaHrlsGlyLyserThrGlnGlnleuLeuSerProLyser 61
QY 718 AAACATTTTCTACTGATAGTGGCGCATGTCACTAACTTGAAGAAAGCTTACTA 777
Db 62 LyserProPheAlaThrAspSerGlyAlaMetProHrlslyslsleuGlnlyslsAspLeu 81
QY 778 AAGGCTAATTGAAGAACAATTGATCGCTAACGTGCACAGTACGACTATTGAGGTC 837
Db 82 LyAlAlalleGlnGlnGlnleuIleAlaAsnValHrlsSerAsnAspArgTyPheGlnVal 101
QY 838 ATTGATTTTGCAGCGCATGCAACATTAATGATCGAAGCAAGCGCAAGTCTACTTGTGAC 897
Db 102 IleAspPheAlaSerAspAlaThrilleThrAspArgAsnGlyrsValTyPheAlaAsp 121
QY 898 AAGAGTGTTCGGGTAACCTTGCCGACCCGACCTGTCCAGAAATTTTGTGAAGCGCAT 957
Db 122 LyAspGlySerValThrleuProHlnProValGlnleuPheleuSerGlyHrls 141
QY 958 GTCGCGGTTAGACCATTAATAAAGAAACCAATCAAAACCAAGCGAAATCTGTGATGTG 1017
Db 142 ValArgValArgProTyrlsGlnlyrsProIleGlnAsnGlnAlaLysserValAspVal 161
QY 1018 GAATATACGTGATACGTTACTCCCTTAAACCTGTATGACATTTACAGACAGTCTCAAA 1077
Db 162 GluTyThrValGlnPheThrProleuAspProAspAspPheArgProGlyLeuLy 181
QY 1078 GATACTAAGCTATTTGAAAACACTAGCTATCGGAGACCACTCAATCTCAAGAAATTA 1137
Db 182 AspThrLyslsleuLeuLyThrleuAlalleGlyAspThrilleThrSerGlnleuLeu 201
QY 1138 GCTCAAGCAAAAGCAATTTTAAACAACCAACCCAGGCTATAGATTTTGAACGTGAC 1197
Db 202 AAGlnAlaGlnserlleuLeuLnlysrThrlsProGlytrThrlleryGluArgAsp 221
QY 1198 TCCTCAATCGTCACTGATGCAATGACATTTTCGATGATTTTACATGATGATCAAGAG 1257
Db 222 SerSerlleValThrHrlsAspAsnspIlePheArgThrilleuProMetAspGlnGlu 241
QY 1258 TTACTTACCGGTAAAAATCGGGAACAAGCTTATAGATCAATAAAAAATCGGTG 1317
Db 242 PheThrlyrHrlsVallyAsnAsnArgGlnGlnAlaTyGlnlleAsnlyslsSerGlyLeu 261
QY 1318 AATGAAGAATTAACAACAACCTGATCTGTGAGAAATATTATCGTCTTAAAAAGG 1377
Db 262 AsnGlnGlnlleAsnAsnThrAspIleSerGlnlyTyTrlyValleuLyslsGly 281
QY 1378 GAAAAGCCGTATGATCCCTTTGATCGCACTGCTTGAACCTGTTCACCAATACGTT 1437
Db 282 GluLyrsProTyrsAspProPheAspArgSerHrlsleuLyslsleuPheThrilleuTyTrVal 301
QY 1438 GATGTCGATACCAAGAAATTCGTAATAAAGTGAGCAGCTTAAACAGTACGAAACGTAC 1497
Db 302 AspValAsnThrAsnGlnleuLeuLyserGlnGlnleuThrAlaSerGlnArgAsn 321
QY 1498 TTAGACTTCAGAGATTTATACGATCCCTGATTAAGGCTAAACTACTTCAACAATCTC 1557
Db 322 LeuAspPheArgAspLeuTyrsAspProArgAspLyAlaLyslsleuLeuTyrsAsnleu 341
QY 1558 GATGCTTTTGGTATTATGACATTAACCTTAACCTGAAAAGTGAAGATTAATCAAGTAC 1617
Db 342 AspAlaPheGlylleMetAspTyThrleuThrGlyValGlnAspAsnHrlsAspAsp 361

QY 1618 ACCAACCGTATCATTAACCGTTATATAGGCGACGCCGAGAGAGAAATGCTAGCTAT 1677
Db 362 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGluValAsnAlaSerTyr 361
QY 1678 CATTTAGCCTATGATTAAGAATGCTTATACCGAAGAAAGAAAGTTTACCTACTG 1737
Db 362 HisLeuAlaIyrAspLysAspArgTyrThrGluGluGluValTyrSerTyrLeu 401
QY 1738 CGTTATACAGGAGACACTTATACCTGATTAACCTTAACGACAA 1779
Db 402 ArgTyrThrGlyThrProIleProAspAsnProAsnAspLys 415
RESULT 5
US-10-300-215-253
; Sequence 253, Application US/10300215
; Publication No. US20030153043A1
; GENERAL INFORMATION:
; APPLICANT: CARR, Francis Joseph
; APPLICANT: ADAIR, Fiona Suzanne
; APPLICANT: HAMILTON, Anita Anne
; APPLICANT: CARTER, Graham
; TITLE OF INVENTION: METHOD FOR THE PRODUCTION OF
; TITLE OF INVENTION: NON-IMMUNOGENIC PROTEINS
; FILE REFERENCE: MER-104-CON.1
; CURRENT APPLICATION NUMBER: US/10/300,215
; PRIOR FILING DATE: 2002-11-20
; PRIOR APPLICATION NUMBER: US 09/438,136
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: WO PCT/GB98/01473
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: GB 9710480.6
; PRIOR FILING DATE: 1997-05-21
; PRIOR APPLICATION NUMBER: GB 9716197.0
; PRIOR FILING DATE: 1997-07-31
; PRIOR APPLICATION NUMBER: GB 9725270.4
; PRIOR FILING DATE: 1997-11-28
; PRIOR APPLICATION NUMBER: GB 9807751.4
; PRIOR FILING DATE: 1998-04-14
; PRIOR APPLICATION NUMBER: US 60/067,235
; PRIOR FILING DATE: 1997-12-02
; NUMBER OF SEQ ID NOS: 254
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 253
; LENGTH: 414
; TYPE: PRP
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Modified strep protein
US-10-300-215-253.
Alignment Scores:
Pred. No.: 9,716-189 Length: 414
Score: 2121.00 Matches: 407
Percent Similarity: 98.79% Conservative: 2
Best Local Similarity: 98.31% Mismatches: 5
Query Match: 67.68% Indels: 0
Gaps: 0
US-09-940-235-11 (1-1782) x US-10-300-215-253 (1-414)
QY 538 ATTGCTGACCGTGAAGTGGCTGTAGCCGTCCTCATCTGTCAACAACGCAATTTGTTT 597
Db 1 IleAlaGlyProGluIrrIrrLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
QY 598 AGCGTTGCTGACTGTTGAGGGGAGCAATCAAGATTAAGTCTTAATTTTGAATC 657
Db 21 SerValAlaGlyThrValGluGluThrAsnGlnAspIleSerLeuLysPhePheGluIle 40
QY 658 GATCTAACATCAAGCACTGCTCATGAGAGAAAGAGAGAGAGAGCTTAAGTCCAAATCA 717
Db 41 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnIlyLeuSerProLysSer 60
QY 718 AAACCATTTGCTACTGATATAGTGGCGGAGATGTCAATTAACCTTGAGAAAGCTGACTTA 777

Db 61 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGluValAspLeuLeu 80
QY 778 AAGGCTATTCAGAAACAATTGATCGCTTACCTCCACATTAACAGACATTAAGTTCAG 837
Db 81 LysAlaLysGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspLysPheGluVal 100
QY 838 ATTGATTTTTCAGACGATGCAACCATTAATCTGATGCAAAACGGCAAGCTTACTTTCAG 897
Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
QY 898 AAAGATGCTTCGTTACCTTCCGACCCCAACCTGTCGAAGATTTTTCAGACGACAT 957
Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
QY 958 GTGGCGCTTACGATCAATTAAGAAACCAATACAAACCAAGCAAGAAATCTGTGATGTG 1017
Db 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
QY 1018 GAATATACCTGATGAGTTTACTCCCTTAAACCTGATGACATTTTCAGACGAGTCCAA 1077
Db 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
QY 1078 GATATCAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACATCTCAAGATTACTA 1137
Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 200
QY 1138 GCTTAAGCAACAAGCAATTTTAAACAAAACCAACCGGCTTATGACATTTATGAACGTGAC 1197
Db 201 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp 220
QY 1198 TCCCTCAATCGCATCTGACCAATGACATTTTCGTCGATTTTACCAATGATCAACAG 1257
Db 221 SerSerIleValIrrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnIu 240
QY 1258 TTTACTTACCGTGTGTTAAAAATCGGAAACAAGCTTATGATCAATTAATAAATCGTGTG 1317
Db 241 PheThrTyrHisValLysValAsnArgGluGlnAlaTyrGluIleAsnLysLysSerGlyLeu 260
QY 1318 AATGAGAAATTAACAACACTGACCTGATCTTGAGAAATATTAAGTCTTAAAAAGG 1377
Db 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280
QY 1378 GAAAAGCGGTATGATCCCTTTGATCGGACGATCTTGAACCTGTTACCATCAATACGTT 1437
Db 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysPheVal 300
QY 1438 GATGTCGATCAACGAATTTGCTAAAAAGTGAAGAGCTTAAACGCTAGGAGAACGTAC 1497
Db 301 AspValAsnThrAsnGlnLeuLeuLysSerGlnGlnLeuLeuThrAlaSerGluArgAsn 320
QY 1498 TTAGACTTCAGAGATTATTAAGATCTCTGATAGGCTTAAGCTTAACATTAACATCTC 1557
Db 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
QY 1558 GATGCTTTTGGTATTAATGACTATACCTTAACCTGAAAAGTGAAGATTAATCAGATGAC 1617
Db 341 AspAlaPheGlyIleMetAspArgTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
QY 1618 ACCAACCGTATCATTAACCGTTATATAGGCGACGCCGAGAGAGAAATGCTAGCTAT 1677
Db 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGluValAsnAlaSerTyr 380
QY 1678 CATTTAGCCTATGATTAAGAATGCTTATACCGAAGAAAGAAAGTTTACCTACTG 1737
Db 381 HisLeuAlaIyrAspLysAspArgTyrThrGluGluGluValArgGluValTyrSerTyrLeu 400
QY 1738 CGTTATACAGGAGACACTTATACCTGATTAACCTTAACGACAA 1779
Db 401 ArgTyrThrGlyThrProIleProAspAsnProAsnAspLys 414
RESULT 6
US-09-919-703-12

```

; Sequence 12, Application US/09919703
; Patent No. US20020165129A1
; GENERAL INFORMATION:
; APPLICANT: Kyrleal, Gerald
; APPLICANT: Rabbitt, Simon W.
; TITLE OF INVENTION: Peptides and Their Use to Ameliorate
; FILE REFERENCE: 50216/003004
; CURRENT APPLICATION NUMBER: US/09/919,703
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 09/294,457
; PRIOR FILING DATE: 1999-04-19
; PRIOR APPLICATION NUMBER: US 08/759,599
; PRIOR FILING DATE: 1996-12-05
; PRIOR APPLICATION NUMBER: US 60/008,233
; PRIOR FILING DATE: 1995-12-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 413
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic polypeptide
; US-09-919-703-12

Alignment Scores:
Pred. No.: 6,84e-186 Length: 413
Score: 2090.50 Matches: 408
Percent Similarity: 98.80% Conservative: 2
Best Local Similarity: 98.31% Mismatches: 2
Query Match: 66.70% Indels: 3
DB: 3 Gaps: 3

US-09-940-235-11 (1-1782) x US-09-919-703-12 (1-413)
QY 538 ATTCGTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAGCAAGCCAAATTGTTGTT 597
DB 1 IleaaglyprrglttrpneuhsparpgrproserValashmsnserglneuVal 20
QY 598 AGCGTGTGCTACTGTGTGAGGAGCAATCAAGCATTAGTCTTAAATTTTGTGAATC 657
DB 21 SerValaIaglyThrVal---GlyThrAsnGlnAspIleSerIeuysrhepnegluile 39
QY 658 GATCTAACATCAAGCACTGCTCATGAGAGAAAGACAGCAAGGCTTAAAGTCCAAATCA 717
DB 40 AspIeuThrSerArgProIahIsglyglyThrGlnGlnIyIeuSerProIysSer 59
QY 718 AAACCATTTGCTACTGATAGTGGCGGATGTCACTAACTTGAGAAAGCTGACTACTA 777
DB 60 LysProPhealathrAspSerGlyalAscSerHislybIeuGlnlysaAspIeuLeu 79
QY 778 AAGCTATTCAAGAACAAATTGATCGCTAACGTCACAGTACAGCACTTATTTGAGGTC 837
DB 80 LysalalIeGlnGlnGlnIeuIeIahAsnValHisSerAsnhsparprrrhegluVal 99
QY 838 ATTGATTTTGCAGCGAGTCAACCATTTACTGTAACGCGCAAGGCTTACTTGTGTAC 897
DB 100 IleAspPhealaserAspAlathrIleThrAspArgAsnGlylysaValIyrrPhealAsp 119
QY 898 AAAGATGTGTGGTAACTTGTGGCGACCCCAACCTGTCACAGAAATTTTGTAGCGGACAT 957
DB 120 LysAspGlySerValThrIeuProThrGlnProValGlnGlnPheIeuSerGlyHis 139
QY 958 GTGCGCGTTAAGACATATTAAGAAAAACCAATACAAACCAAGCGAAATCTGTGATGTC 1017
DB 140 ValaIagValaIargProIyrrIyIeSgluysrProIleGlnAsnGlnIahIyaserValaAsp 159
QY 1018 GAATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTCAGACCAAGGTCGAA 1077
DB 160 GluIyrrThrValaGlnPheThrProIeuAsnProAspAspAspAspAspAspAspAsp 179
QY 1078 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACATCACTCATCTCAAGATTACTA 1137

```

```

DB 180 LeuThrIySleuIeuysrThrIeAlaIleGlyAspThrIleThrSerGlnIeuLeu 199
QY 1138 GCTCAAGCACAAAGCATTTTAAACAAAACCAACCAAGGCTTACGATTTATGAAACGGTAC 1197
DB 200 AlaGlnAlaGlnSerIleIeuAsnlyAsnIleProIyrrThrIleIyrrGluArgAsp 219
QY 1198 TCCCATGCTGCACTACATGACATGACATTTCCGTCGATTTTACCAAGATCAAGAG 1257
DB 220 SerSerIleValThrHisAspAsnAspIlePheArgThrIleIeuProIeuAspGlnGlu 239
QY 1258 TTTACTTACCGTGTAAAAATCGGAAACAAGCTTATGATCAATATAAAAAATCGTCTG 1317
DB 240 PheThrIyrrValGlyAlaIyAsnAsnArgGlnGlnAlaIyrrArgIleAsnlySlySerGlyLeu 259
QY 1318 AATGAAGAAATTAACAACACTGACCTGATCTTCGAGAAAT--TACGTCCTTAAAAA 1374
DB 260 AsnGlnGlnIuIleAsnAsnThrAspIeuIleSerIeuGlnIyrrIyrrValIeuIySly 279
QY 1375 GGGGAAAAAGCCGATGATCCCTTGATCGGACGCTTGAACCTGACCAATCAATAC 1434
DB 280 GlyIuysrProIyrrAspProPheAspArgSerHisIeuysrIeuPheThrIleIyrr 299
QY 1435 GTTGATGCTCGATACCAACGAATGCTTAAAAAGTACGAGGCTTAAACAGCTAGCGAAGT 1494
DB 300 ValaAspValaAspThrAsnGlnIeuIeuIySerGlnIeuIeuThrAlaSerGluArg 319
QY 1495 AACTTACACTTCAGACATTTATACATCTCTGTGATACGCTTAACTTACATCAAT 1554
DB 320 AsnIeuAspPheArgAspIeuIyrrAspProIyrrAspIyrrAlaIySlyIeuIyrrAsn 339
QY 1555 CTGCAAGCTTTTGGTATTATGACCTATCCTTAACTGGAAGAGTAAAGATCAAGAT 1614
DB 340 LeuAspAlaPheGlyIleIeuAspIyrrThrIeuThrGlyIyrrValGluAspAsnHisAsp 359
QY 1615 GACACCAACCGTTCATTAACGTTTATATGAGGACGCAACCGAAGAGAGAAATGCTAGC 1674
DB 360 AspThrAsnAspGlyIleIeuThrValIyrrMetGlyIyrrArgProGlnGlnIuIyAsnHisSer 379
QY 1675 TATCATTTAGCTTATGATTAAGATCGTTATACGGAAGACAGAGAGTTTACAGTAC 1734
DB 380 TyrHis---AlaIyrrAspIyrrAspArgIyrrThrGlnGlnIuIyrrGluValIyrrSerIyrr 398
QY 1735 CTGCGTTTATACAGGAGCACTTATACCTGATTAACCTTACAGCAAA 1779
DB 399 LeuArgIyrrThrGlyThrProIleProAspAsnProAspAspIyrr 413

RESULT 7
US-10-474-792-658
; Sequence 658, Application US/10474792
; Publication No. US20040236072A1
; GENERAL INFORMATION:
; APPLICANT: Olmsted, Stephen
; APPLICANT: Zagursky, Robert
; APPLICANT: Nickbarg, Elliot
; APPLICANT: Winter, Louie
; TITLE OF INVENTION: SURFACE PROTEINS OF STREPTOCOCCUS PYOGENES
; FILE REFERENCE: AM 100399
; CURRENT APPLICATION NUMBER: US/10/474,792
; NUMBER OF SEQ ID NOS: 674
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 658
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes
; US-10-474-792-658

Alignment Scores:
Pred. No.: 1.51e-171 Length: 440
Score: 1937.00 Matches: 367
Percent Similarity: 93.48% Conservative: 20
Best Local Similarity: 88.65% Mismatches: 27

```


Query Match:	61.81%	Indels:
DB:	5	Gaps:
US-09-940-235-11 (1-1782) x US-10-474-792-658 (1-440)		

[illegible]

Oy	1558	GATCGTTTGGTATTTAAGACTATACCTTAACGTGAAGAGGATTAATCAACGATGAC	1617
Db	367	AspAlaIleAspIleMetAspTyrThrLeuThrIlyValGIuAspAsnHtAspLys	386
Oy	1618	ACCAACCGTATCATTAACCGTTTATATGGGCAAGCCGCCGAGAGAGAAATGCTAGCTAT	1677
Db	387	AsnAsnArgValIaThrValThrValTyrMetGlyLysArgProLysGlyAlaLysGlySerTyr	406
Oy	1678	CATTAGCCCTTATGATTAAGATCGTTATACGAGAGAGAGAGAGATTTCAGCTACCTG	1737
Db	407	HisLeuAlaTyrAspLysAspLeuTyrThrGIuGIuGIuIuArgLysAlaTyrSerTyrLeu	426
Oy	1738	CGTATATACAGGACACCATTAACCTGATACCCCTTAAGACAAA	1779
Db	427	ArgAspThrIleGlyThrProIleProAspAsnProLysAspLys	440
RESULT 8			
US-10-236-392-2			
Sequence 2, Application US/10236392			
Publication No. US20040067490A1			
GENERAL INFORMATION:			
APPLICANT: Anderson, David W			
APPLICANT: Boldog, Ferenc L			
APPLICANT: Burgess, Catherine, E			
APPLICANT: Caeman, Stacie J			
APPLICANT: Carterton, Elina			
APPLICANT: Chapoval, Andrei			
APPLICANT: Crabtree, Julie			
APPLICANT: Edinger, Shlomit, R			
APPLICANT: Ellerman, Karen			
APPLICANT: Gerlach, Valerie			
APPLICANT: Gorman, Linda			
APPLICANT: Groose, William M			
APPLICANT: Gusev, Vladimír			
APPLICANT: Kekuda, Ramesh			
APPLICANT: LaRocheHelle, William J			
Li, Li			
APPLICANT: MacDougall, John R			
APPLICANT: Malyankar, Uriel M			
APPLICANT: Miller, Charles E			
APPLICANT: Miller, Isabelle			
APPLICANT: Padigaru, Muralidhara			
APPLICANT: Patturajan, Meera			
APPLICANT: Pena, Carol A			
APPLICANT: Peyman, John A			
APPLICANT: Rastelli, Luca			
APPLICANT: Reiger, Daniel K			
APPLICANT: Rothenberg, Mark E			
APPLICANT: Shenoy, Suresh			
APPLICANT: Shinkete, Richard A			
APPLICANT: Sultison, Glenda			
TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME			
FILE REFERENCE: 21402-442A			
CURRENT APPLICATION NUMBER: US/10/236,392			
CURRENT FILING DATE: 2002-09-06			
PRIOR APPLICATION NUMBER: US09/540,763			
PRIOR FILING DATE: 2000-03-30			
PRIOR APPLICATION NUMBER: US60/390,155			
PRIOR FILING DATE: 2002-06-19			
PRIOR APPLICATION NUMBER: US09/635,949			
PRIOR FILING DATE: 2000-08-10			
PRIOR APPLICATION NUMBER: US60/318,765			
PRIOR FILING DATE: 2001-09-12			
PRIOR APPLICATION NUMBER: US60/357,303			
PRIOR FILING DATE: 2002-02-15			
PRIOR APPLICATION NUMBER: US60/367,753			
PRIOR FILING DATE: 2002-03-25			
PRIOR APPLICATION NUMBER: US60/369,479			
PRIOR FILING DATE: 2002-04-02			
PRIOR APPLICATION NUMBER: US09/659,634			
PRIOR FILING DATE: 2000-09-12			
PRIOR APPLICATION NUMBER: US60/318,120			
PRIOR FILING DATE: 2001-09-07			

```

; PRIOR APPLICATION NUMBER: US60/318,130
; PRIOR FILING DATE: 2001-09-07
; Remaining Prior Application data removed - See File Wrapper or PAM.
; NUMBER OF SEQ ID NOS: 794
; SOFTWARE: Custom
; SEQ ID NO 2
; LENGTH: 2320
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-236-392-2

Alignment Scores:
Pred. No.: 1,666-37 Length: 2320
Score: 504.50 Matches: 105
Percent Similarity: 60.10% Conservative: 11
Best Local Similarity: 54.40% Mismatches: 28
Query Match: 16.10% Indels: 50
DB: Gaps: 4

US-09-940-235-11 (1-1782) x US-10-236-392-2 (1-2320)

QY 254 CCCATAGCTGAGAGTGTGTTGATCATGCTGCTGGAGCTTCTATGTGTCGAGAAACG 313
Db 146 Proilealaglulycyspheasphialaiaaglythrserlyvalaiglylthr
QY 314 TGGGA-----GAGGAGC 327
Db 166 TrrglulysprotyrclnglytrpmetmetvalasrcysthrCysleuglyglulysr 185
QY 328 GAGCGCATCTTGACCTTCTAGAAATAGATGCAACATCAGACACAGACATCTCTAT 387
Db 186 GlyaagliehrCysThrserArghanaArgCysasnsrghlnsrhrthrserlyr 205
QY 388 AGAATTGAGACACCTGAGCAAGAGATTAATCGAGAACTGCTCGATGCAATGTC 447
Db 206 ArgileglyasprthrtrpserlylsasrpsasndrglyAsnleuenglnCysleCys 225
QY 448 ACAGCGCAACGCGGAGAGAGAGTGGAAGTGTGAGAGGACACCTCTGTGACACCATCG 507
Db 226 ThrGlyasnglyArgrgllyutrrplysCysgluarghlsrthrserValglnthrThrser 245
QY 508 AGCGGATCTGGCCCTTCACCGATGTCGATTGCT----- 543
Db 246 SerGlyserGlyProphethrAspValaArgAlaIaValIyrglnProglnProhlspro 265
QY 544 -----GACCT 549
Db 266 GlnProProtyrGlyHisCysValThrAspserGlyValIyrservalGlymet 285
QY 550 GAGTGGCTGTAGACCGTCATCTGTCAACAACAGCCAAATGGTTGACGTTGCTGCT 609
Db 286 Glntrpleu-----LysThrGlnGlyAsnlysglnmetleuCystrCysleugly 302
QY 610 ACTGTGAGGGGAGCAATCAAGACATTAAGTCTTAATTTTGAATTCAGATCAACATCA 669
Db 303 AsnglyValserCysglnGlnutrrAlaValThrGlnThrTyrelYclysAsnleuAnGly 322
QY 670 GACCT-----GCTCATGAGAGAAAGCA 693
Db 323 GluProCysValleuProphethrTyrrAsnglyArgrThr 335

RESULT 9
US-10-171-311-64
; Sequence 64, Application US/10171311
; Publication No. US20030087270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Glatc, Karen
; APPLICANT: Gannavarapu, Manjula

```

```

; APPLICANT: Hoereth, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 64
; LENGTH: 2328
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-171-311-64

Alignment Scores:
Pred. No.: 1,666-37 Length: 2328
Score: 504.50 Matches: 105
Percent Similarity: 60.10% Conservative: 11
Best Local Similarity: 54.40% Mismatches: 28
Query Match: 16.10% Indels: 50
DB: Gaps: 4

US-09-940-235-11 (1-1782) x US-10-171-311-64 (1-2328)

QY 254 CCCATAGCTGAGAGTGTGTTGATCATGCTGCTGGAGCTTCTATGTGTCGAGAAACG 313
Db 154 Proilealaglulycyspheasphialaiaaglythrserlyvalaiglylthr
QY 314 TGGGA-----GAGGAGC 327
Db 174 TrrglulysprotyrclnglytrpmetmetvalasrcysthrCysleuglyglulysr 193
QY 328 GAGCGCATCTTGACCTTCTAGAAATAGATGCAACATCAGACACAGACATCTCTAT 387
Db 194 GlyaagliehrCysThrserArghanaArgCysasnsrghlnsrhrthrserlyr 213
QY 388 AGAATTGAGACACCTGAGCAAGAGATTAATCGAGAACTGCTCGATGCAATGTC 447
Db 214 ArgileglyasprthrtrpserlylsasrpsasndrglyAsnleuenglnCysleCys 233
QY 448 ACAGCGCAACGCGGAGAGAGTGGAAGTGTGAGAGGACACCTCTGTGACACCATCG 507
Db 234 ThrGlyasnglyArgrgllyutrrplysCysgluarghlsrthrserValglnthrThrser 253
QY 508 AGCGGATCTGGCCCTTCACCGATGTCGATTGCT----- 543
Db 254 SerGlyserGlyProphethrAspValaArgAlaIaValIyrglnProglnProhlspro 273
QY 544 -----GACCT 549
Db 274 GlnProProtyrGlyHisCysValThrAspserGlyValIyrservalGlymet 293
QY 550 GAGTGGCTGTAGACCGTCATCTGTCAACAACAGCCAAATGGTTGACGTTGCTGCT 609
Db 294 Glntrpleu-----LysThrGlnGlyAsnlysglnmetleuCystrCysleugly 310
QY 610 ACTGTGAGGGGAGCAATCAAGACATTAAGTCTTAATTTTGAATTCAGATCAACATCA 669
Db 311 AsnglyValserCysglnGlnutrrAlaValThrGlnThrTyrelYclysAsnleuAnGly 330
QY 670 GACCT-----GCTCATGAGAGAAAGCA 693
Db 331 GluProCysValleuProphethrTyrrAsnglyArgrThr 343

RESULT 10
US-10-236-031B-70

```

```
/ Sequence 70, Application US/10236031B
/ Publication No. US20030219760A1
/ GENERAL INFORMATION:
/ APPLICANT: Gordon, Gavin J.
/ APPLICANT: Jensen, Rodrick V.
/ APPLICANT: Gullans, Steven R.
/ APPLICANT: Bueno, Raphael
/ TITLE OF INVENTION: Diagnostic and Prognostic Tests
/ FILE REFERENCE: B00801/70265 (JRV/JAV)
/ CURRENT APPLICATION NUMBER: US/10/236, 031B
/ CURRENT FILING DATE: 2002-09-05
/ PRIOR APPLICATION NUMBER: US 60/317,389
/ PRIOR FILING DATE: 2001-09-05
/ PRIOR APPLICATION NUMBER: US 60/407,431
/ PRIOR FILING DATE: 2002-08-30
/ NUMBER OF SEQ ID NOS: 102
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 70
/ LENGTH: 2328
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-236-031B-70
```

```
Alignment Scores:
Pred. No.: 1.66e-37 Length: 2328
Score: 504.50 Matches: 105
Percent Similarity: 60.10% Conservative: 11
Best Local Similarity: 54.40% Mismatches: 28
Query Match: 16.10% Indels: 50
DB: Gaps: 4
```

```
US-09-940-235-11 (1-1782) x US-10-236-031B-70 (1-2328)
QY 254 CCCATGCTGGAAGGTTTGTGATCATGCTGCGGACCTTCCTATGCTGGAGAAACG 313
DB 154 ProllealaglulysCyspheasphlsalalaglYThrserYValValGlyGlnThr 173
QY 314 TGGGA-----GAAGGAGC 327
DB 174 TTPGluYsPrOtyrGlnGlyTTPMeMetValAspCysThrCybLeuGlyGlnGlySer 193
QY 328 GGAAGCATCACTTGCATTGTAGAAATAGATGCAACGATGAGACACAAAGACATCTTAT 387
DB 194 GLyArgIleThrCystrSerArgAsnArgCybAsnAspGlnAspThrArgThrSerYr 213
QY 388 AGAATTGAGACACCTGAGACAAAGAAAGATATCGAGAAACCTGCTCCAGTGCAATGCG 447
DB 214 ArgIleGlyAspThrTPserYsLysAspAsnArgGlyAsnLeuLeuGlnCybIleCys 233
QY 448 ACAAGCAAGGCGGAGAGAGAGTGAAGTGAAGGACACACTCTGTGACAGACATCG 507
DB 234 ThrGlyAsnGlyArgGlyGlnTTPYsCybGlnArgHisThrSerValGlnThrThrSer 253
QY 508 AGCGAGTGGCCCTTCAACCGATGTTCCGATTGCT-----543
DB 254 SerGlySerGlyPProPheThrAspValArgAlaValYrGlnProGlnProHisPro 273
QY 544 -----GAGACT 549
DB 274 GlnProProProtyrGlyHisCysValThrAspSerGlyValValYrSerValGlyMet 293
QY 550 GAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCAATGCGTTGTTAGCGTTGCTGCT 609
DB 294 GlnTPLeu-----LysThrGlnGlyAsnLysGlnMetLeuCybThrCybLeuGly 310
QY 610 ACTGTTGAGGGGAGCAATGACATTAAGCTTAATTTTGAATTCGATCAATCAATCA 669
DB 311 AsnGlyValSerCybGlnGlnThrAlaValThrGlnThrYrGlyGlyAsnLeuAsnGly 330
QY 670 CGACCT-----GCTCATGGAGGAAGACA 693
DB 331 GlnProCybValLeuProPheThrYrAsnGlyArgThr 343
```

```
RESULT 11
US-10-374-979-98
/ Sequence 98, Application US/10374979
/ Publication No. US20030219793A1
/ GENERAL INFORMATION:
/ APPLICANT: John P. Carulli et al.
/ TITLE OF INVENTION: THE HIGH BONE MASS GENE OF 11q13.3
/ FILE REFERENCE: 032796-021
/ CURRENT APPLICATION NUMBER: US/10/374, 979
/ CURRENT FILING DATE: 2003-03-04
/ PRIOR APPLICATION NUMBER: US 09/544, 398
/ PRIOR FILING DATE: 2000-04-05
/ PRIOR APPLICATION NUMBER: US 09/543, 771
/ PRIOR FILING DATE: 2000-04-05
/ PRIOR APPLICATION NUMBER: US 09/229, 319
/ PRIOR FILING DATE: 1999-01-13
/ PRIOR APPLICATION NUMBER: US 60/071, 449
/ PRIOR FILING DATE: 1998-01-13
/ PRIOR APPLICATION NUMBER: US 60/105, 511
/ PRIOR FILING DATE: 1998-10-23
/ NUMBER OF SEQ ID NOS: 109
/ SEQ ID NO 98
/ LENGTH: 2328
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-374-979-98
```

```
Alignment Scores:
Pred. No.: 1.66e-37 Length: 2328
Score: 504.50 Matches: 105
Percent Similarity: 60.10% Conservative: 11
Best Local Similarity: 54.40% Mismatches: 28
Query Match: 16.10% Indels: 50
DB: Gaps: 4
```

```
US-09-940-235-11 (1-1782) x US-10-374-979-98 (1-2328)
QY 254 CCCATGCTGGAAGGTTTGTGATCATGCTGCGGACCTTCCTATGCTGGAGAAACG 313
DB 154 ProllealaglulysCyspheasphlsalalaglYThrserYValValGlyGlnThr 173
QY 314 TGGGA-----GAAGGAGC 327
DB 174 TTPGluYsPrOtyrGlnGlyTTPMeMetValAspCysThrCybLeuGlyGlnGlySer 193
QY 328 GGAAGCATCACTTGCATTGTAGAAATAGATGCAACGATGAGACACAAAGACATCTTAT 387
DB 194 GLyArgIleThrCystrSerArgAsnArgCybAsnAspGlnAspThrArgThrSerYr 213
QY 388 AGAATTGAGACACCTGAGACAAAGAAAGATATCGAGAAACCTGCTCCAGTGCAATGCG 447
DB 214 ArgIleGlyAspThrTPserYsLysAspAsnArgGlyAsnLeuLeuGlnCybIleCys 233
QY 448 ACAAGCAAGGCGGAGAGAGTGAAGTGAAGGACACACTCTGTGACAGACATCG 507
DB 234 ThrGlyAsnGlyArgGlyGlnTTPYsCybGlnArgHisThrSerValGlnThrThrSer 253
QY 508 AGCGAGTGGCCCTTCAACCGATGTTCCGATTGCT-----543
DB 254 SerGlySerGlyPProPheThrAspValArgAlaValYrGlnProGlnProHisPro 273
QY 544 -----GAGACT 549
DB 274 GlnProProProtyrGlyHisCysValThrAspSerGlyValValYrSerValGlyMet 293
QY 550 GAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCAATGCGTTGTTAGCGTTGCTGCT 609
DB 294 GlnTPLeu-----LysThrGlnGlyAsnLysGlnMetLeuCybThrCybLeuGly 310
QY 610 ACTGTTGAGGGGAGCAATGACATTAAGCTTAATTTTGAATTCGATCAATCAATCA 669
DB 311 AsnGlyValSerCybGlnGlnThrAlaValThrGlnThrYrGlyGlyAsnLeuAsnGly 330
```

```

QY 670 CGACCT-----GCTCATGAGGAAAGACA 693
Db 331 GluProCysValLeuProPheThrTyrAsnGlyArgThr 343

RESULT 12
US-10-182-936A-98
; Sequence 98, Application US/10182936A
; Publication No. US20040038860A1
GENERAL INFORMATION:
; APPLICANT: Allen, Kristina M.
; APPLICANT: Anisowicz, Anthony
; APPLICANT: Bhat, Bheem
; APPLICANT: Damagnez, Veronique
; APPLICANT: Robinson, John
; APPLICANT: Yaworsky, Paul
TITLE OF INVENTION: Reagents and Method for Modulating DKK-Mediated Interactions
FILE REFERENCE: 032796-143
CURRENT APPLICATION NUMBER: US/10/182,936A
PRIOR FILING DATE: 2002-08-02
PRIOR APPLICATION NUMBER: PCT/US02/15982
PRIOR FILING DATE: 2002-05-17
PRIOR APPLICATION NUMBER: US 60/291,311
PRIOR FILING DATE: 2001-05-17
PRIOR APPLICATION NUMBER: US 60/353,058
PRIOR FILING DATE: 2002-02-01
PRIOR APPLICATION NUMBER: US 60/361,293
PRIOR FILING DATE: 2002-03-04
NUMBER OF SEQ ID NOS: 216
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 98
LENGTH: 2328
TYPE: PRT
ORGANISM: Homo sapiens
US-10-182-936A-98

Alignment Scores:
Pred. No.: 1,666-37 Length: 2328
Score: 504.50 Matches: 105
Percent Similarity: 60.10% Conservative: 11
Best Local Similarity: 54.40% Mismatches: 28
Query Match: 16.10% Indels: 50
DB: 4 Gaps: 4

US-09-940-235-11 (1-1782) x US-10-182-936A-98 (1-2328)
QY 254 CCCATAGCTGAGAAAGTTTGTATGATCATGCTGGCACTTCTATGTGTCGAGAAACG 313
Db 154 ProilealagluCyserPheaspHisalaalaglyThrSerValValGlyGluThr 173
QY 314 TGGGA-----GAGGCAGC 327
Db 174 TrpGluysProTyrgInglyTrpMetMetValAspCysThrCysleuGlyGluGlySer 193
QY 328 GAGCGATCACTTGCACTTCTAGAAATAGATGCAACGATCAGAGACAAGCATCTAT 387
Db 194 GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr 213
QY 388 AGAATTGAGACACCTGAGCAAGAAGATTAATCGAAGAACTGCTCCAGTCGATCTGC 447
Db 214 ArgIleGlyAspThrTrpSerTyrSlyAspAsnArgGlyAsnleuGluGlySlyCys 233
QY 448 ACAGCGAACCGCCGAGAGAGTGGAAGTGTGAGGCGACACTCTGTGCAACCATCG 507
Db 234 ThrGlyAsnGlyArgGlyGlyTrpTyrSlySlyGluArgHisThrSerValGlnThrThrSer 253
QY 508 AGCGGATCGGCCCTTACCGATGTCGATTGCT-----GAGCCT 543
Db 254 SerGlySerGlyProPheThrAspValArgAlaAlaValTyrGlnProGlnProHisPro 273
QY 544 -----GAGCCT 549
Db 274 GlnProProTyrglyHisCysValThrAspSerGlyValValTyrSerValGlyMet 293

```

```

QY 550 GAGTGGCTGTAGACCGTCATCTGTCAACAAGCCAAATGGGTGTACGTTGCTGT 609
Db 294 GlnTrpLeu-----LysThrGlnGlyAsnLysGlnMetLeuCystrThrCysleuGly 310
QY 610 ACTGTTGAGGGAGCAATCAAGACATTAGTCTTAATTTTGTGAATCGATCAACATCA 669
Db 311 AsnGlyValSerCysGlnGluThrAlaValThrGlnThrTyrGlyGlyAsnleuAsnGly 330
QY 670 CGACCT-----GCTCATGAGGAAAGACA 693
Db 331 GluProCysValLeuProPheThrTyrAsnGlyArgThr 343

RESULT 13
US-10-477-238A-677
; Sequence 677, Application US/10477238A
; Publication No. US20040221326A1
GENERAL INFORMATION:
; APPLICANT: Babi, Philip
; APPLICANT: Yaworsky, Paul
; APPLICANT: Bex, Frederick J. III
; APPLICANT: Bodine, Peter Van Nest
TITLE OF INVENTION: Transgenic Animal Model of Bone Mass Modulation
FILE REFERENCE: 032796-212
CURRENT APPLICATION NUMBER: US/10/477,238A
PRIOR FILING DATE: 2003-11-10
PRIOR APPLICATION NUMBER: US 60/290,071
PRIOR FILING DATE: 2001-05-11
PRIOR APPLICATION NUMBER: US 60/291,311
PRIOR FILING DATE: 2001-05-17
PRIOR APPLICATION NUMBER: US 60/353,058
PRIOR FILING DATE: 2002-02-01
PRIOR APPLICATION NUMBER: US 60/361,293
PRIOR FILING DATE: 2002-03-04
NUMBER OF SEQ ID NOS: 812
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 677
LENGTH: 2328
TYPE: PRT
ORGANISM: Homo sapiens
US-10-477-238A-677

Alignment Scores:
Pred. No.: 1,666-37 Length: 2328
Score: 504.50 Matches: 105
Percent Similarity: 60.10% Conservative: 11
Best Local Similarity: 54.40% Mismatches: 28
Query Match: 16.10% Indels: 50
DB: 5 Gaps: 4

US-09-940-235-11 (1-1782) x US-10-477-238A-677 (1-2328)
QY 254 CCCATAGCTGAGAAAGTTTGTATGATCATGCTGGCACTTCTATGTGTCGAGAAACG 313
Db 154 ProilealagluCyserPheaspHisalaalaglyThrSerValValGlyGluThr 173
QY 314 TGGGA-----GAGGCAGC 327
Db 174 TrpGluysProTyrgInglyTrpMetMetValAspCysThrCysleuGlyGluGlySer 193
QY 328 GAGCGATCACTTGCACTTCTAGAAATAGATGCAACGATCAGAGACAAGCATCTAT 387
Db 194 GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr 213
QY 388 AGAATTGAGACACCTGAGCAAGAAGATTAATCGAAGAACTGCTCCAGTCGATCTGC 447
Db 214 ArgIleGlyAspThrTrpSerTyrSlyAspAsnArgGlyAsnleuGluGlySlyCys 233
QY 448 ACAGCGAACCGCCGAGAGAGTGGAAGTGTGAGGCGACACTCTGTGCAACCATCG 507
Db 234 ThrGlyAsnGlyArgGlyGlyTrpTyrSlySlyGluArgHisThrSerValGlnThrThrSer 253
QY 508 AGCGGATCGGCCCTTACCGATGTCGATTGCT-----GAGCCT 543

```

```
Db 254 SerGlySerGlyProPheThrAspValArgAlaIaValTyrGlnProGlnProHisPro 273
Qy 544 -----GACCT 549
Db 274 GlnProProTyrGlyHisCysValThrAspSerGlyValValTyrSerValGlyMet 293
Qy 550 GAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCAATTGGTTGGTTCGTTGCTGT 609
Db 294 GlnTrpLeu-----LysThrGlnGlyAsnLysGlnMetLeuCysThrCysLeuGly 310
Qy 610 ACTGTTGAGGGAGCAAGCATCAAGACATTAGCTTAATTTTGAATCCGATCTAACATCA 669
Db 311 AsnGlyValSerCysGlnGlnThrAlaValThrGlnThrTyrGlyGlyAsnLeuAsnGly 330
Qy 670 CGACCT-----GCTCATGAGAGAAAGACA 693
Db 331 GluProCysValLeuProPheThrTyrAsnGlyArgThr 343

RESULT 14
US-10-680-287A-677
; Sequence 677, Application US/10680287A
; Publication No. US20040244069A1
; GENERAL INFORMATION:
; APPLICANT: BabiJ, Philip
; APPLICANT: Yaworsky, Paul
; APPLICANT: Bex, Frederick J. III
; APPLICANT: Bodine, Peter Van Nest
; TITLE OF INVENTION: Transgenic Animal Model of Bone Mass Modulation
; FILE REFERENCE: 032796-179
; CURRENT APPLICATION NUMBER: US/10/680,287A
; CURRENT FILING DATE: 2003-10-08
; PRIOR APPLICATION NUMBER: PCT/US02/14876
; PRIOR FILING DATE: 2002-05-13
; PRIOR APPLICATION NUMBER: US 60/290,071
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/291,311
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: US 60/353,058
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: US 60/361,293
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 812
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 677
; LENGTH: 2328
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-680-287A-677

Alignment Scores:
Pred. No.: 1,666-37 Length: 2328
Score: 504.50 Matches: 105
Percent Similarity: 60.10% Conservative: 11
Best Local Similarity: 54.40% Mismatches: 28
Query Match: 16.10% Indels: 50
DB: 5 Gaps: 4

US-09-940-235-11 (1-1782) x US-10-680-287A-677 (1-2328)
Qy 254 CCCATAGCTGAGAGTGTTTTGTATGCTGTGGACTTCCTATGNGTGGAGAAAG 313
Db 154 ProtlehlaGluLysCysPheAspHisAlaIaGlyThrSerTyrValIaGlyLutThr 173
Qy 314 TGGGA-----GAAGCAGC 327
Db 174 TrpGluLysProTyrGlnGlyTTrpMetMetValAspCysThrCysLeuGlyGluGlySer 193
Qy 328 GGAAGCATCTTTCGACTTCTAGAAATAGATGCAAGCATGAGACAACAGACATCTTAT 387
Db 194 GlyArgGlyLeuThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr 213
Qy 388 AGAATTGGAGACACCTGGAGCAAGAGATTAATCGAGAAACTGCTCCAGTGCATCTGC 447
```

```
Db 214 ArgIleGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuGlnCysIleCys 233
Qy 448 ACAGCAACGGCCGAGAGAGTGTGAGAGCAACCTGTGTGACACCAATCG 507
Db 234 ThrGlyAsnGlyArgGlyGlnTrpLysCysGluArgHisThrSerValGlnThrSer 253
Qy 508 AGCGATCTGAGCCCTTCACCGAATGTTGCTATTTGCT----- 543
Db 254 SerGlySerGlyProPheThrAspValArgAlaIaValTyrGlnProGlnProHisPro 273
Qy 544 -----GACCT 549
Db 274 GlnProProTyrGlyHisCysValThrAspSerGlyValValTyrSerValGlyMet 293
Qy 550 GAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCAATTGGTTGGTTCGTTGCTGT 609
Db 294 GlnTrpLeu-----LysThrGlnGlyAsnLysGlnMetLeuCysThrCysLeuGly 310
Qy 610 ACTGTTGAGGGAGCAAGCATCAAGACATTAGCTTAATTTTGAATCCGATCTAACATCA 669
Db 311 AsnGlyValSerCysGlnGlnThrAlaValThrGlnThrTyrGlyGlyAsnLeuAsnGly 330
Qy 670 CGACCT-----GCTCATGAGAGAAAGACA 693
Db 331 GluProCysValLeuProPheThrTyrAsnGlyArgThr 343

RESULT 15
US-10-477-173-677
; Sequence 677, Application US/10477173
; Publication No. US20050070699A1
; GENERAL INFORMATION:
; APPLICANT: Genome Therapeutics Corporation and
; APPLICANT: Allen, Kristina M.
; APPLICANT: Yaworsky, Paul
; APPLICANT: Morales, Arturo J.
; APPLICANT: Graham, James R.
; APPLICANT: Anisowicz, Anthony
; APPLICANT: Liu, Wei
; TITLE OF INVENTION: HBM Variants that Modulate Bone Mass and Lipid Levels
; FILE REFERENCE: 032796-135
; CURRENT APPLICATION NUMBER: US/10/477,173
; CURRENT FILING DATE: 2003-11-10
; PRIOR APPLICATION NUMBER: US 60/290,071
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/291,311
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: US 60/353,058
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: US 60/361,293
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 1086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 677
; LENGTH: 2328
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-477-173-677

Alignment Scores:
Pred. No.: 1,666-37 Length: 2328
Score: 504.50 Matches: 105
Percent Similarity: 60.10% Conservative: 11
Best Local Similarity: 54.40% Mismatches: 28
Query Match: 16.10% Indels: 50
DB: 5 Gaps: 4

US-09-940-235-11 (1-1782) x US-10-477-173-677 (1-2328)
Qy 254 CCCATAGCTGAGAGTGTTTTGTATGCTGTGGACTTCCTATGNGTGGAGAAAG 313
Db 154 ProtlehlaGluLysCysPheAspHisAlaIaGlyThrSerTyrValIaGlyLutThr 173
Qy 314 TGGGA-----GAAGCAGC 327
```

```
Db      174 TrpGluLysProTyrGlnGlyTTrpMetMetValAspCysThrCysLeuGlyGluGlySer 193
      |||
QY      328 GGACGCATCATCTTGCACTTCTAGAAATATGATGCAACGATCAGACACACAGACATCTCTAT 387
      |||
Db      194 GLyArgGLeThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr 213
      |||
QY      388 AGAATTGAGAGACACCTGAGCAAGAGATTAATCGAGAAACCTGCTCCAGTCATCTGC 447
      |||
Db      214 ArgGLeGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuLeuGlnCys1LeCys 233
      |||
QY      448 ACAGGCAACGCGCGAGAGAGATGGAAGTGTGAGAGGCACACCTCTGTGACAGACACATCG 507
      |||
Db      234 ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSer 253
      |||
QY      508 AGCGATCTGGCCCTTCAACCGATGTCGTATTGCT----- 543
      |||
Db      254 SerGlySerGlyProPheThrAspValArgAlaAlaValTyrGlnProGlnProHisPro 273
      |||
QY      544 -----GACCT 549
      |||
Db      274 GlnProProTyrGlyHisCysValThrAspSerGlyValValTyrSerValGlyMet 293
      |||
QY      550 GAGTGGCTGCTAGACCGTCATCTGTCAACAACAGCCAAATGSGTGTAGCTTGCTGCT 609
      |||
Db      294 GlnTrpLeu-----LysThrGlnGlyAsnLysGlnMetLeuLeuCysThrCysLeuGly 310
      |||
QY      610 ACTGTTGAGGGGAGCAGATCAAGACATTAAGTCTTAATTTTGAATTCGATCTACATCA 669
      |||
Db      311 AsnGlyValSerCysGlnGluThrAlaValThrGlnThrTyrGlyGlyAsnLeuAsnGly 330
      |||
QY      670 CGACCT-----GCTCATGAGAGAAAGCA 693
      |||
Db      331 GlnProCysValLeuProPheThrTyrAsnGlyArgThr 343
      |||
```

Search completed: January 28, 2006, 02:57:14
Job time : 140.535 secs

Pred. No.: 6.3e-157 Length: 414
 Score: 2158.00 Matches: 414
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 68.86% Indels: 0
 DB: 6 Gaps: 0

US-09-940-235-11 (1-1782) x US-10-631-558-2 (1-414)

QY	538	ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAAATTGGTTGT	597
DB	1	llealaglProglutlrpleuLeuaspArgProSerValasnbnserGlnleuVal	20
QY	598	AGCGTTGCGTACTGTTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTGAATC	657
DB	21	ServalAlaaglYthrValgluGlnYthrAsnGlnAspIleSerleuYsPhepegIuile	40
QY	658	GATCTAACATCAAGACCTGCTCATGAGGAAAGACAGACAGGCTTAAGTCCAAATCA	717
DB	41	AspleuThrSerArgProAlaHlsGlyGlyYsThrGlnGlnIlyleuSerProlySer	60
QY	718	AAACCATTTGCTACTGATATGTCGCGCATGTCACATAACTTGAGAAAGCTGACTTACTA	777
DB	61	LysProPheAlaThrAspSerGlyAlaMetSerHlsYleuGlnIlySAlaAspleuLeu	80
QY	778	AAGCTATTCAAGACAATGATTCGCTACGTCACAGTACAGACGACTTACTTGAGCTC	837
DB	81	LysAlaAlleGlnGlnGlnleuIleAlaAsnValHlsSerAsnAspAspYrPheGlnVal	100
QY	838	ATTGATTTTGCAGCGATGCAACCATTAATCTGATCGAAAGCGCAAGTCTACTTGTCTGAC	897
DB	101	IleAspPheAlaSerAspAlaThrIleThrAspAspAsnGlyYsValYrPheAlaAsp	120
QY	898	AAAGATGTTGCGTAACCTTGGCCGACCCGACCTGTCCAGAAATTTTGTAAAGCGACAT	957
DB	121	LysAspGlySerValThrleuProThrGlnProValGlnIlyPheleuLeuSerGlyHls	140
QY	958	GTCGCGCTTACAGCATATPAAGAAAAACCAATCAAAAACCAAGCCGAATCTGTATGTG	1017
DB	141	ValArgValArgProYrIlySgIlySProIleGlnAsnGlnAlaYsSerValaAspVal	160
QY	1018	GAATATACGTGACGTTTACTCCCTTAAACCTGATGACGATTTGAGACAGGTCCAA	1077
DB	161	GluYrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuYs	180
QY	1078	GATACATAAGCTATTGAAAACTAGCTATCGGTGACACCATCAATCTCAAGATTACTA	1137
DB	181	AspThrYsleuLeuYsThrleuAlaIleGlyAspThrIleThrSerGlnIlyleuLeu	200
QY	1138	GCTCAAGCAAGCAATTTTAAACAAAACCCAGGCTATACGATTTATGACGTGAC	1197
DB	201	AlaGlnAlaGlnSerIleleuAsnYsAsnHlsProGlyYrThrIleYrGluArgAsp	220
QY	1198	TCTCAATCGTCACTCATGACATGATTTCCGATGATTTTACCATGATGATCAAG	1257
DB	221	SerSerIleValIleHlsAspAsnAspIlePheAspGlnThrIleuProMetSerGlnGlu	240
QY	1258	TTTACTTACCGTGTAAATAATCGGAAACAAGCTTATAGATCAATAAATACTGGTCTG	1317
DB	241	PheThrYrArgValIlyAsnAsnArgIuGlnAlaYrArgIleAsnYsYsSerGlyLeu	260
QY	1318	AATGAGAAATPAACAACATGACCTGATCTCTGAGAAATATTAGCTCTTAAAAAAGG	1377
DB	261	AsnGlnGlnIleAsnAsnThrAspLeuIleSerGlnYrYrYrValleuYsYsGly	280
QY	1378	GAAGAAGCGTATGATCCCTTGAATGCGAGTCACTGGAACCTGTCCACCAATTCGTT	1437
DB	281	GluYsProYrAspProPheAspArgSerHlsleuYsleuPheThrIleYsYrVal	300
QY	1438	GATGTCATACCAAGAAATTTGCTAAAAAGTGACAGCTCTTAAAGCTACCAACGTAC	1497
DB	301	AspValaAspThrAsnGlnleuLeuYsSerGlnGlnleuLeuThrHlsSerGlnArgAsn	320

QY	1498	TTAGACTTCAGAGATTTATAGATTCCTGATTAAGGCTAAACTACTCAACATCTC	1557
DB	321	LeuAspPheArgAspleuYrAspProAlaAspYsAlaYsleuLeuYrAsnAsnleu	340
QY	1558	GATGCTTTTGGTATTATGACATATACCTTAACCTGAAAGAGATGATTAACAGATGAC	1617
DB	341	AspAlaPheGlyIleMetAspYrThrleuThrGlyYsValGlnAspAsnHlsAspAsp	360
QY	1618	ACCAACCGTATCATTAACCGTTTATATGGCGACAGACCCGAAAGAGAAATGCTACTAT	1677
DB	361	ThrAsnArgIleIleThrValYrMetGlyYsArgProGlnGlnIlyGlnAsnAlaSerYr	380
QY	1678	CATTAGCCTATGATTAAGATCGTTATATACCGAAGAAACAAGAAAGTTTACGACTCTG	1737
DB	381	HlsleuAlaYrAspYsAspArgYrThrGlnGlnIlyGlnArgGlnValYrSerYrleu	400
QY	1738	CGTTATACAGGACACCTATACCTGATTAACCTTAACGACAA	1779
DB	401	ArgYrThrGlyThrProIleProAsnProAsnAspYs	414

RESULT 2
 US-11-032-951-12
 ; Sequence 12, Application US/11032951
 ; Publication No. US20050282741A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kyteal, Gerald
 ; APPLICANT: Rabkin, Simon W.
 ; TITLE OF INVENTION: Peptides and Their Use to Ameliorate
 ; FILE REFERENCE: 50216/003005
 ; CURRENT APPLICATION NUMBER: US/11/032,951
 ; PRIOR FILING DATE: 2005-01-11
 ; PRIOR APPLICATION NUMBER: US 09/919,703
 ; PRIOR FILING DATE: 2001-07-31
 ; PRIOR APPLICATION NUMBER: US 09/294,457
 ; PRIOR FILING DATE: 1999-04-19
 ; PRIOR APPLICATION NUMBER: US 08/759,599
 ; PRIOR FILING DATE: 1996-12-05
 ; PRIOR APPLICATION NUMBER: US 60/008,233
 ; PRIOR FILING DATE: 1995-12-06
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 12
 ; LENGTH: 413
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Synthetic polypeptide
 US-11-032-951-12

Alignment Scores:
 Pred. No.: 8.47e-152 Length: 413
 Score: 2090.50 Matches: 408
 Percent Similarity: 98.80% Conservative: 2
 Best Local Similarity: 98.31% Mismatches: 2
 Query Match: 66.70% Indels: 3
 DB: 7 Gaps: 3

US-09-940-235-11 (1-1782) x US-11-032-951-12 (1-413)

QY	538	ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAAATTGGTTGT	597
DB	1	llealaglProglutlrpleuLeuaspArgProSerValasnbnserGlnleuVal	20
QY	598	AGCGTTGCGTACTGTTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTGAATC	657
DB	21	ServalAlaaglYthrVal--GlyThrAsnGlnAspIleSerleuYsPhepegIuile	39
QY	658	GATCTAACATCAAGACCTGCTCATGAGGAAAGACAGACAGGCTTAAGTCCAAATCA	717
DB	40	AspleuThrSerArgProAlaHlsGlyGlyYsThrGlnGlnIlyleuSerProlySer	59
QY	718	AAACCATTTGCTACTGATATGTCGCGCATGTCACATAACTTGAGAAAGCTGACTTACTA	777


```

Db 60 LysProPheAlaThrAspSerGlyAlaMetSerHisLeuLeuGlyAlaAspLeuLeu 79
QY 778 AAGGCTATTCAAGAACAAATGATCGCTTAAGCTCAAGCAAGCTGTTGAGATC 837
Db 80 LysAlaAlaLeuGlnGlnLeuLeuLeuAlaAsnValHisSerAsnAspArgTyrPheGlnVal 99
QY 838 ATTGATTTTGCAGCGCATGCAACCATTAATGATGCAAAAGCGCAAGTCTTCTTGCTGAC 897
Db 100 IleAspPheAlaSerThrAlaThrIleThrAspArgAsnIlyValTyrPheAlaAsp 119
QY 898 AAGATGTTCCGTTAACTTGGCCGACCCAACTGTTCCAGAAATTTTGTCTAAGCGGACAT 957
Db 120 LysAspGlySerValThrLeuProThrGlnProValGlnGlnIlePheLeuLeuSerGlyHis 139
QY 958 GTGGCGGTTAGAACCATATTAAGAAAACCAATACCAACGCGAAATCTGTGATGATG 1017
Db 140 ValArgValArgProTyrLysGlnLysProIleGlnAsnGlnAlaLysSerValAspVal 159
QY 1018 GAATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTCAAGCAGGCTCAAA 1077
Db 160 GluTyrThrValGlnThrThrProLeuAsnProAspAspAspPheArgProGlyLeuLys 179
QY 1078 GATPACTAAGCTATTGAAAACACTAGCTATCGGTGACACATCAATCTCAAGAAATTA 1137
Db 180 LeuThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeuLeu 199
QY 1138 GCTCAAGCAACAAAGCAATTTTAAACAAACCAACCGGCTATACGATTTTGAACGTCAC 1197
Db 200 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGlnLysArg 219
QY 1198 TCTCGATGCTGCTGATGACATGATTTTCCGATGATTTTCCAAAGGATGACAG 1257
Db 220 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGln 239
QY 1258 TTTACTTACCGTGTAAATAATCGGAAACAGCTTATAGATCAATTAATAATCGTGTG 1317
Db 240 PheThrTyrArgValLysAsnArgGlnGlnAlaIleTyrArgIleAsnLysLysSerGlyLeu 259
QY 1318 AATGAAGAAATAAACCAACACTGATCTGTGAGAAATAT--TACGCTTTAAAAA 1374
Db 260 AsnGlnGlnIleLeuAsnThrAspLeuIleSerLeuGlnTyrLysTyrValLeuLys 279
QY 1375 GGGGAAAACCGGATATGCTTGTGATCGAGTCACTGAAACTGTTACACATCAATAC 1434
Db 280 GlyLysLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyr 299
QY 1435 GTTGAATGTCATACCAAGAAATGCTAAAAAAGTGAAGAGCTTTAACAGCTAGGAAAGT 1494
Db 300 ValAspValAspThrAsnGlnLeuLeuLysSerGlnGlnLeuThrAlaSerGlnArg 319
QY 1495 AACTTGAAGCTTCAAGATTTATACGATCTCGTATAGAGCTTAACTACTTCAACAAAT 1554
Db 320 AsnLeuAspPheArgSerLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsn 339
QY 1555 CTGCGATGCTTTTGGTATTTATGACTATACCTTATCTGAAAGAGATTAATCACGAT 1614
Db 340 LeuAspAlaPheGlyIleMetAspArgTyrThrLeuThrGlyLysValGlnAspAsnHisAsp 359
QY 1615 GACACCAACCGTATCAACCGTTTATATGAGGACAGACCGAAGAGAGAAATGCTAGC 1674
Db 360 AspThrAsnArgIleIleThrValTyrMetGlyLysArgProGlnGlnLysValAsnAlaSer 379
QY 1675 TATCATTTTACCTATGATTAAGATCTTATACCGAAGAGAGAAAGTTCACGATC 1734
Db 380 TyrHis--AlaTyrAspLysAspArgTyrThrGlnGlnGlnLysValIlyrSerTyr 398
QY 1735 CTGCGTTATACAGGAGACCTTATACCTGATACCTTAAACGACAA 1779
Db 399 LeuArgTyrThrGlyThrProIleProAspAsnProAspAspLys 413

```

RESULT 3
US-10-995-561-631

```

; Sequence 631, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 631
; LENGTH: 642
; TYPE: PR
; ORGANISM: Homo sapiens
US-10-995-561-631

Alignment Scores:
Pred. No.: 3,77e-31 Length: 642
Score: 503.00 Matches: 101
Percent Similarity: 66.25% Conservative: 5
Best Local Similarity: 63.12% Mismatches: 10
Query Match: 16.05% Indels: 45
DB: 6 Gaps: 3

US-09-940-235-11 (1-1782) x US-10-995-561-631 (1-642)
QY 254 CCCATGACCTGAGAAAGTGTGATCATGCTGCTGAGACTTCTATGTTGTCGAGAAACG 313
Db 181 ProIleAlaGlnLysCysPheAspHisAlaIleGlyThrSerTyrValValGlyGlnThr 200
QY 314 TGGGA-----GAAAGCAGC 327
Db 201 TrpGlnLysProTyrGlnGlnTyrTrpMetValAspCysThrCysLeuGlyGlnGlySer 220
QY 328 GAGCGCATCACTTGACCTTGTAGAAATAGATGACATGACAGACAGCAAGCAATCCAT 387
Db 221 GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr 240
QY 388 AGAATTGAGACACCTGAGCAAGAGATGATGAGAAACCTGCTTCAAGTCAATGCTGC 447
Db 241 ArgIleGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuLeuGlnCysIleCys 260
QY 448 ACAGCGAACGCGGAGAGAGTGAAGTGTGAGAGGCAACCTGTGTGACAGCAATCG 507
Db 261 ThrGlyAsnGlyArgGlyGlnTyrLysCysGlnArgHisThrSerValGlnThrSer 280
QY 508 AGCGGATCTGGCCCTTCCACCGATGTCGATTGCT-----543
Db 281 SerGlySerGlyProPheThrAspValArgAlaValTyrGlnProGlnProHisBro 300
QY 544 -----GGAAGCT 549
Db 301 GlnProProProTyrGlyHisCysValThrAspSerGlyValValTyrSerValGlyMet 320
QY 550 GAGTGGCTGCTAGACCGTCACTCTGTCAACAACGCAATGAGTGTGTTAGCGTGTGCT 609
Db 321 GlnTrpLeu-----LysThrGlnGlyAsnLysGlnMetLeuCysThrCysLeuGly 337

RESULT 4
US-10-995-561-622
; Sequence 622, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702

```

SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO: 622
 LENGTH: 657
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-995-561-622

Alignment Scores:
 Pred. No.: 3,79e-31 Length: 657
 Score: 503.00 Matches: 101
 Percent Similarity: 66.25% Conservative: 5
 Best Local Similarity: 63.12% Mismatches: 10
 Query Match: 16.05% Indels: 45
 Gaps: 3

US-09-940-235-11 (1-1782) x US-10-995-561-622 (1-657)

```

QY 254 CCATAGCTGAGAGAGTGTGGATGATGCTGGAGCTTCTATGTGTGGAGAAACG 313
DB 181 ProilealaglulycyspheaphisalaialaglyThrseryValValglulThr 200
QY 314 TGGGA-----GAGGCAGC 327
DB 201 TrpGlulysProtyrGlnGlyTrpMetMetValAspCysThrCysleuGlyGlnGlySer 220
QY 328 GAGCAGCATCTTGACATTTAGAAATAGATGCAAGATCAGGACAGACATCTTAT 387
DB 221 GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr 240
QY 388 AGAATTGAGAGACACTGGAGCAAGAGATTAATCGAAGAACTGCTCCAGTCATCTGC 447
DB 241 ArgIleGlyAspThrTrpSerIlyAspAsnArgGlyAsnleuGlnCysIleCys 260
QY 448 ACAAGCAACGCGCCGAGAGAGAGTGAAGTGTAGAGGCACACTCTGTGACACCATCG 507
DB 261 ThrGlyAsnGlyArgGlyGlyIleTrpIlyCysGlyArgHisThrSerValGlnThrThrSer 280
QY 508 AGCGGATCTGGCCCTTACCGATGTCGTAATGCT-----GAGCCT 543
DB 281 SerGlySerGlyProphethrAspValArgAlaAlaValTyrGlnProGlnProHisPro 300
QY 544 -----GAGCCT 549
DB 301 GlnProProProTyrglyHisCysValThrAspSerGlyValValTyrSerValGlyMet 320
QY 550 GAGTGGCTGCTAGACCGCTCATCTGTCAACAACGCAATGTGGTGTAGCGTTGCTGCT 609
DB 321 GlnTrpLeu-----LysThrGlnGlyAsnIlyGlnMetleuCysThrCysleuGly 337

```

RESULT 5

US-10-995-561-629
 Sequence 629, Application US/10995561
 Publication No. US20050272054A1
 GENERAL INFORMATION:
 APPLICANT: CARGILL, Michele et al.
 TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
 FILE REFERENCE: CL001559
 CURRENT APPLICATION NUMBER: US/10/995,561
 CURRENT FILING DATE: 2004-11-24
 NUMBER OF SEQ ID NOS: 85702
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO: 629
 LENGTH: 984
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-995-561-629

Alignment Scores:
 Pred. No.: 4.05e-31 Length: 984
 Score: 503.00 Matches: 101
 Percent Similarity: 66.25% Conservative: 5

Best Local Similarity: 63.12% Mismatches: 10
 Query Match: 16.05% Indels: 45
 DB: 6 Gaps: 3

US-09-940-235-11 (1-1782) x US-10-995-561-629 (1-984)

```

QY 254 CCATAGCTGAGAGAGTGTGGATGATGCTGGAGCTTCTATGTGTGGAGAAACG 313
DB 181 ProilealaglulycyspheaphisalaialaglyThrseryValValglulThr 200
QY 314 TGGGA-----GAGGCAGC 327
DB 201 TrpGlulysProtyrGlnGlyTrpMetMetValAspCysThrCysleuGlyGlnGlySer 220
QY 328 GAGCAGCATCTTGACATTTAGAAATAGATGCAAGATCAGGACAGACATCTTAT 387
DB 221 GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr 240
QY 388 AGAATTGAGAGACACTGGAGCAAGAGATTAATCGAAGAACTGCTCCAGTCATCTGC 447
DB 241 ArgIleGlyAspThrTrpSerIlyAspAsnArgGlyAsnleuGlnCysIleCys 260
QY 448 ACAAGCAACGCGCCGAGAGAGAGTGAAGTGTAGAGGCACACTCTGTGACACCATCG 507
DB 261 ThrGlyAsnGlyArgGlyGlyIleTrpIlyCysGlyArgHisThrSerValGlnThrThrSer 280
QY 508 AGCGGATCTGGCCCTTACCGATGTCGTAATGCT-----GAGCCT 543
DB 281 SerGlySerGlyProphethrAspValArgAlaAlaValTyrGlnProGlnProHisPro 300
QY 544 -----GAGCCT 549
DB 301 GlnProProProTyrglyHisCysValThrAspSerGlyValValTyrSerValGlyMet 320
QY 550 GAGTGGCTGCTAGACCGCTCATCTGTCAACAACGCAATGTGGTGTAGCGTTGCTGCT 609
DB 321 GlnTrpLeu-----LysThrGlnGlyAsnIlyGlnMetleuCysThrCysleuGly 337

```

RESULT 6

US-10-995-561-633
 Sequence 633, Application US/10995561
 Publication No. US20050272054A1
 GENERAL INFORMATION:
 APPLICANT: CARGILL, Michele et al.
 TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
 FILE REFERENCE: CL001559
 CURRENT APPLICATION NUMBER: US/10/995,561
 CURRENT FILING DATE: 2004-11-24
 NUMBER OF SEQ ID NOS: 85702
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO: 633
 LENGTH: 2296
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-995-561-633

Alignment Scores:
 Pred. No.: 4.67e-31 Length: 2296
 Score: 503.00 Matches: 101
 Percent Similarity: 66.25% Conservative: 5
 Best Local Similarity: 63.12% Mismatches: 10
 Query Match: 16.05% Indels: 45
 DB: 6 Gaps: 3

US-09-940-235-11 (1-1782) x US-10-995-561-633 (1-2296)

```

QY 254 CCATAGCTGAGAGAGTGTGGATGATGCTGGAGCTTCTATGTGTGGAGAAACG 313
DB 181 ProilealaglulycyspheaphisalaialaglyThrseryValValglulThr 200
QY 314 TGGGA-----GAGGCAGC 327

```

```
Db      201  TTPGluYsProTyrgInglYTrpMeCMeValAspCySthrCySleuGlYglUglYser 220
OY      328  GGAGCATGCTGCTGCACTTGTAGAAATAGATGCAAGCATGAGACAGACATCTCAT 387
Db      221  GYAAGGileThrCySthrsrArghsAnArgCySbAnspGlnAspThrArghThrsrTy 240
OY      388  AGAATTGAGACACCTGAGACAGAAAGATATCGAAGAACTGCTCCATGTCATGTC 447
Db      241  ArgIleGlYAspThrTrpSerIlySlyAspAsnArgIlyAsnLeuLeuGlnCySleCys 260
OY      448  ACAAGGACGCGCCGAGAGAGAGTGAAGTGTGAGAGCACACTTGTGACAGACCATCG 507
Db      261  ThrGlyAsnGlyArGslYglUtrPlySlySgluArghIsthSrValGlnThrThrsr 280
OY      508  AGCGGATCGGCGCCCTTCAACCGATGTCGTAATGCT----- 543
Db      281  SerGlySerGlyProPheThrAspValArgAlaIalValTyrgInProGlnProHisPro 300
OY      544  -----GGACT 549
Db      301  GlnProProProTyrgIlyHisCySValThrAspSerGlyValValTyrsrValGlyMet 320
OY      550  GAGTGGCTGCTAGACCGCTCATCTGTCAACAACAGCCAAATGGTTGTTAGCGTTCGT 609
Db      321  GlnTrpLeu-----LysThrGlnGlyAsnIlySglInMetLeuCySthrCySleuGlY 337
```

RESULT 7

```
US-10-995-561-623
; Sequence 623, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 623
; LENGTH: 2355
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-623
```

Alignment Scores:

```
Pred. No.: 4,69e-31 Length: 2355
Score: 503.00 Matches: 101
Percent Similarity: 66.25% Conservative: 5
Best Local Similarity: 63.12% Mismatches: 10
Query Match: 16.05% Indels: 45
DB: 6 Gaps: 3
```

US-09-940-235-11 (1-1782) x US-10-995-561-623 (1-2355)

```
OY      254  CCCATAGCTGAGAAAGTGTGATGATGCTGTGGAGCTTCTATGTGTGAGAAACG 313
Db      181  ProIleAlaGluYsCySphneAspHisAlaIalGlyThrsrTyValValGlyGluThr 200
OY      314  TGGGA-----GAAAGCAGC 327
Db      201  TrpGluYsProTyrgInglYTrpMeCMeValAspCySthrCySleuGlYglUglYser 220
OY      328  GGAGCATGCTGCTGCACTTGTAGAAATAGATGCAAGCATGAGACAGACATCTCAT 387
Db      221  GYAAGGileThrCySthrsrArghsAnArgCySbAnspGlnAspThrArghThrsrTy 240
OY      388  AGAATTGAGACACCTGAGACAGAAAGATATCGAAGAACTGCTCCATGTCATGTC 447
Db      241  ArgIleGlYAspThrTrpSerIlySlyAspAsnArgIlyAsnLeuLeuGlnCySleCys 260
```

```
OY      448  ACAAGGACGCGCCGAGAGAGTGAAGTGTGAGAGCACACTTGTGACAGACCATCG 507
Db      261  ThrGlyAsnGlyArGslYglUtrPlySlySgluArghIsthSrValGlnThrThrsr 280
OY      508  AGCGGATCGGCGCCCTTCAACCGATGTCGTAATGCT----- 543
Db      281  SerGlySerGlyProPheThrAspValArgAlaIalValTyrgInProGlnProHisPro 300
OY      544  -----GGACT 549
Db      301  GlnProProProTyrgIlyHisCySValThrAspSerGlyValValTyrsrValGlyMet 320
OY      550  GAGTGGCTGCTAGACCGCTCATCTGTCAACAACAGCCAAATGGTTGTTAGCGTTCGT 609
Db      321  GlnTrpLeu-----LysThrGlnGlyAsnIlySglInMetLeuCySthrCySleuGlY 337
```

RESULT 8

```
US-10-995-561-627
; Sequence 627, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 627
; LENGTH: 2355
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-627
```

Alignment Scores:

```
Pred. No.: 4,69e-31 Length: 2355
Score: 503.00 Matches: 101
Percent Similarity: 66.25% Conservative: 5
Best Local Similarity: 63.12% Mismatches: 10
Query Match: 16.05% Indels: 45
DB: 6 Gaps: 3
```

US-09-940-235-11 (1-1782) x US-10-995-561-627 (1-2355)

```
OY      254  CCCATAGCTGAGAAAGTGTGATGATGCTGTGGAGCTTCTATGTGTGAGAAACG 313
Db      181  ProIleAlaGluYsCySphneAspHisAlaIalGlyThrsrTyValValGlyGluThr 200
OY      314  TGGGA-----GAAAGCAGC 327
Db      201  TrpGluYsProTyrgInglYTrpMeCMeValAspCySthrCySleuGlYglUglYser 220
OY      328  GGAGCATGCTGCTGCACTTGTAGAAATAGATGCAAGCATGAGACAGACATCTCAT 387
Db      221  GYAAGGileThrCySthrsrArghsAnArgCySbAnspGlnAspThrArghThrsrTy 240
OY      388  AGAATTGAGACACCTGAGACAGAAAGATATCGAAGAACTGCTCCATGTCATGTC 447
Db      241  ArgIleGlYAspThrTrpSerIlySlyAspAsnArgIlyAsnLeuLeuGlnCySleCys 260
OY      448  ACAAGGACGCGCCGAGAGAGTGAAGTGTGAGAGCACACTTGTGACAGACCATCG 507
Db      261  ThrGlyAsnGlyArGslYglUtrPlySlySgluArghIsthSrValGlnThrThrsr 280
OY      508  AGCGGATCGGCGCCCTTCAACCGATGTCGTAATGCT----- 543
Db      281  SerGlySerGlyProPheThrAspValArgAlaIalValTyrgInProGlnProHisPro 300
OY      544  -----GGACT 549
Db      301  GlnProProProTyrgIlyHisCySValThrAspSerGlyValValTyrsrValGlyMet 320
```

```

QY 550 GAGTGGCTGCTAGACCGTCATCTGTCAACAAGCCCAATGTTGTTAGCGTTCGAT 609
Db 321 GlnTrpLeu-----LysThrGlnGlyAsnLysGlnMetLeuCythrCysLeuGly 337

RESULT 9
US-10-821-234-1545
; Sequence 1545, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; PRIOR FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_Seq_genes Version 1.0
; SEQ ID NO 1545
; LENGTH: 2384
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1545

Alignment Scores:
Pred. No.: 4.7e-31 Length: 2384
Score: 503.00 Matches: 101
Percent Similarity: 66.25% Conservative: 5
Best Local Similarity: 63.12% Mismatches: 10
Query Match: 16.05% Indels: 45
DB: Gaps: 3

US-09-940-235-11 (1-1782) x US-10-821-234-1545 (1-2384)
QY 254 CCCATGCTGAGAAAGTGTGTTGATCATGCTGCGAATCTTCTATGTGTCGAGAAAG 313
Db 210 ProtleaGlnLysCysPheAspHisAlaAlaGlyThrSerTyValValGlyGlnThr 229
QY 314 TGGGA-----GAAAGCAGC 327
Db 230 TrpGlnLysProTyGlnGlnGlyTrpMetMetValAspCysThrCysLeuGlyGlnGlySer 249
QY 328 GAGCGATCACTTGCACTTCTAGAAATAGATGCAACGATCAGACACAAGACATCTTAT 387
Db 250 GlyArgGlyLeuThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTy 269
QY 388 AGAATTGAGACACCTGAGACAAAGATTAATCGAAGAACTGCTCCAGTCGATCTGC 447
Db 270 ArgIleGlyAspThrTrpSerTyLysAspAsnArgGlyAsnLeuLeuGlnCysIleCys 289
QY 448 ACAGGCAACGGCCGAGAGAGATGTAAGTGTGAGAGGACACCTCTGTGAGACACATG 507
Db 290 ThrGlyAsnGlyAspGlyGlnTrpLysCysGlnAsnGlnThrSerValGlnThrThrSer 309
QY 508 AGCGGATCGGCCCTTCAACCGATGTCCTATTGCT----- 543
Db 310 SerGlySerGlyProPheThrAspValArgAlaValTyGlnProGlnProHisPro 329
QY 544 -----GAACT 549
Db 330 GlnProProProTyGlyHisCysValThrAspSerGlyValValTySerValGlyMet 349
QY 550 GAGTGGCTGCTAGACCGTCATCTGTCAACAAGCCCAATGTTGTTAGCGTTCGAT 609
Db 350 GlnTrpLeu-----LysThrGlnGlyAsnLysGlnMetLeuCythrCysLeuGly 366

RESULT 10
US-10-995-561-626
; Sequence 626, Application US/10995561

```

```

; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 626
; LENGTH: 2386
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-626

Alignment Scores:
Pred. No.: 4.7e-31 Length: 2386
Score: 503.00 Matches: 101
Percent Similarity: 66.25% Conservative: 5
Best Local Similarity: 63.12% Mismatches: 10
Query Match: 16.05% Indels: 45
DB: Gaps: 3

US-09-940-235-11 (1-1782) x US-10-995-561-626 (1-2386)
QY 254 CCCATGCTGAGAAAGTGTGTTGATCATGCTGCGAATCTTCTATGTGTCGAGAAAG 313
Db 181 ProtleaGlnLysCysPheAspHisAlaAlaGlyThrSerTyValValGlyGlnThr 200
QY 314 TGGGA-----GAAAGCAGC 327
Db 201 TrpGlnLysProTyGlnGlnGlyTrpMetMetValAspCysThrCysLeuGlyGlnGlySer 220
QY 328 GAGCGATCACTTGCACTTCTAGAAATAGATGCAACGATCAGACACAAGACATCTTAT 387
Db 221 GlyArgGlyLeuThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTy 240
QY 388 AGAATTGAGACACCTGAGACAAAGATTAATCGAAGAACTGCTCCAGTCGATCTGC 447
Db 241 ArgIleGlyAspThrTrpSerTyLysAspAsnArgGlyAsnLeuLeuGlnCysIleCys 260
QY 448 ACAGGCAACGGCCGAGAGAGTGTGTAAGTGTGAGAGGACACCTCTGTGCAACACATG 507
Db 261 ThrGlyAsnGlyAspGlyGlnTrpLysCysGlnAsnGlnThrSerValGlnThrThrSer 280
QY 508 AGCGGATCGGCCCTTCAACCGATGTCCTATTGCT----- 543
Db 281 SerGlySerGlyProPheThrAspValArgAlaValTyGlnProGlnProHisPro 300
QY 544 -----GAACT 549
Db 301 GlnProProProTyGlyHisCysValThrAspSerGlyValValTySerValGlyMet 320
QY 550 GAGTGGCTGCTAGACCGTCATCTGTCAACAAGCCCAATGTTGTTAGCGTTCGAT 609
Db 321 GlnTrpLeu-----LysThrGlnGlyAsnLysGlnMetLeuCythrCysLeuGly 337

RESULT 11
US-11-006-119-31
; Sequence 31, Application US/11006119
; Publication No. US20050260691A1
; GENERAL INFORMATION:
; APPLICANT: Ndao, Momar
; APPLICANT: Ward, Brian
; APPLICANT: Caffrey, Rebecca
; APPLICANT: Spithill, Terry
; APPLICANT: Li, Hongshan
; APPLICANT: Podust, Vladimir
; APPLICANT: Perichon, Regis
; APPLICANT: Ciplergen Biosystems, Inc.
; TITLE OF INVENTION: Serum Biomarkers for Chagas Disease

```

FILE REFERENCE: 016866-012130US
CURRENT APPLICATION NUMBER: US/11/006,119
CURRENT FILING DATE: 2004-12-06
PRIOR APPLICATION NUMBER: US 60/527,153
PRIOR FILING DATE: 2003-12-05
PRIOR APPLICATION NUMBER: US 60/565,093
PRIOR FILING DATE: 2004-04-22
PRIOR APPLICATION NUMBER: US 60/625,519
PRIOR FILING DATE: 2004-11-06
NUMBER OF SEQ ID NOS: 38
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 31
LENGTH: 300
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: N-terminal 28.7 fragment of Fibronectin
US-11-006-119-31

Alignment Scores:
Pred. No.: 5,14e-31 Length: 300
Score: 500.50 Matches: 95
Percent Similarity: 84.82% Conservative: 0
Best Local Similarity: 84.82% Mismatches: 2
Query Match: 15.97% Indels: 16
DB: 7 Gaps: 1

US-09-940-235-11 (1-1782) x US-11-006-119-31 (1-300)

QY 254 CCATAGCTGAGAAAGTGTGATCATGCTGCGGACCTTCTATGTGTGGAGAAACG 313
|||
DB 150 Protlealaglulwscyspheaerphialalaglythrserlyvalvalglulthr 169
QY 314 TGGGA-----GAGGCACG 327
|||
DB 170 TrpdlulwspProtyrginglyTrpMetMetValaerCysThrCysleuglyglulysr 189
QY 328 GGATGCATCTGACATCTTCTAGAAATAGATGCAACGATGAGACACATCTCTAT 387
|||
DB 190 GlyArgilethrCysThrSerArgasnArgCysasnberglinspThrArgThrSer 209
QY 388 AGAATTGAGACACCTGAGCAAGAAAGATATGAGAAACCTGCTCAAGTCATCTGC 447
|||
DB 210 ArgileglyAspThrTrpserlyslsAspAsnArglyAsnleuGlnCyslleCys 229
QY 448 AAGCGAAGCGCCGAGAGAGAGTGTGAGAGGACACCTCTGTGACACCATCG 507
|||
DB 230 ThrGlyasnGlyArglyglultrpLysCysgluArgHisThrSerValGlnThrThrser 249

QY 508 AGCGATCTGGCCCTTCAACCGATGTCGATTGCT 543
|||
DB 250 SerGlySerGlyProPheThrAspValArgAlaAla 261

RESULT 12
US-10-631-558-4
Sequence 4, Application US/10631558
Publication No. US20050260598A1
GENERAL INFORMATION:
APPLICANT: Kumar, Rajesh
APPLICANT: Sahni, Girish
APPLICANT: Roy, Chait
APPLICANT: Rajagopal, Kamnara
APPLICANT: Nihalani, Deepak
APPLICANT: Sundaram, Vasudha
APPLICANT: Yadav, Mahavir
TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
TITLE OF INVENTION: PROTEIN
FILE REFERENCE: 07064-009002
CURRENT APPLICATION NUMBER: US/10/631,558
CURRENT FILING DATE: 2003-07-31
PRIOR APPLICATION NUMBER: US/09/940,235

PRIOR FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: 09/471,349
PRIOR FILING DATE: 1999-12-23
PRIOR APPLICATION NUMBER: IN 3825/DEL/98
PRIOR FILING DATE: 1998-12-24
NUMBER OF SEQ ID NOS: 28
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 259
TYPE: PRT
ORGANISM: Homo sapiens
US-10-631-558-4

Alignment Scores:
Pred. No.: 8,48e-31 Length: 259
Score: 497.50 Matches: 94
Percent Similarity: 85.45% Conservative: 0
Best Local Similarity: 85.45% Mismatches: 1
Query Match: 15.87% Indels: 16
DB: 6 Gaps: 1

US-09-940-235-11 (1-1782) x US-10-631-558-4 (1-259)

QY 254 CCATAGCTGAGAAAGTGTGATCATGCTGCGGACCTTCTATGTGTGGAGAAACG 313
|||
DB 150 Protlealaglulwscyspheaerphialalaglythrserlyvalvalglulthr 169
QY 314 TGGGA-----GAGGCACG 327
|||
DB 170 TrpdlulwspProtyrginglyTrpMetMetValaerCysThrCysleuglyglulysr 189
QY 328 GGATGCATCTGACATCTTCTAGAAATAGATGCAACGATGAGACACATCTCTAT 387
|||
DB 190 GlyArgilethrCysThrSerArgasnArgCysasnberglinspThrArgThrSer 209
QY 388 AGAATTGAGACACCTGAGCAAGAAAGATATGAGAAACCTGCTCAAGTCATCTGC 447
|||
DB 210 ArgileglyAspThrTrpserlyslsAspAsnArglyAsnleuGlnCyslleCys 229
QY 448 AAGCGAAGCGCCGAGAGAGAGTGTGAGAGGACACCTCTGTGACACCATCG 507
|||
DB 230 ThrGlyasnGlyArglyglultrpLysCysgluArgHisThrSerValGlnThrThrser 249

QY 508 AGCGATCTGGCCCTTCAACCGATGTCGATTGCT 537
|||
DB 250 SerGlySerGlyProPheThrAspValArg 259

RESULT 13
US-10-995-561-632
Sequence 632, Application US/10995561
Publication No. US20050272054A1
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
TITLE OF INVENTION: DETECTION AND USES THEREOF
FILE REFERENCE: C1001559
CURRENT APPLICATION NUMBER: US/10/995,561
CURRENT FILING DATE: 2004-11-24
NUMBER OF SEQ ID NOS: 85702
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 632
LENGTH: 693
TYPE: PRT
ORGANISM: Homo sapiens
US-10-995-561-632

Alignment Scores:
Pred. No.: 1,41e-05 Length: 693
Score: 166.50 Matches: 36
Percent Similarity: 44.83% Conservative: 16
Best Local Similarity: 31.03% Mismatches: 41
Query Match: 5.31% Indels: 24

```

DB:                               6          Gaps:          4
US-09-940-235-11 (1-1782) x US-10-995-561-632 (1-693)
QY 254 CCCATAGCTGAGAAAGTTTGTATCATGCTGAGCACTTCTATGTGTGAGAAACG 313
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 508 ProthraspaserCyspHeasprProTyThrValSerHisTyrAlaValGlyAspGlu 527
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 314 TGGAGAA-----GGCAGC 327
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 528 TrpGluArgMetSerGluSerGlyPheLeuLeuGlyGlnCysGlnCysLeuGlyPheGlySer 547
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 328 GAGCGCATCTGCACTTCTAGAAATATGATGCATGACATGACAGACACATCCTAT 387
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 548 GlyHisPheArgCysAspSerSerArgTyrCysHisAspAsnGly-----ValAsnTyr 565
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 388 AGAATTGAGACACCTGAGACAGAAAGATATATGAGAAACCTGCTCCAGTCATCTGC 447
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 566 LysIleGlyGlyLysTrpAspArgGlnGlyGluAsnGlyGlnMetMetSerCysThrCys 585
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 448 ACAAGCAACGCCGAGAGAGATGGAAGTGTGAGAGGACACCTCTGTGACACCAATCG 507
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 586 LeuGlyAsnGlyLysGlyGlnPheLysCysAspProHis-----GluAlaThrCys 602
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 508 AGCGGATCTGGCCCTTCAACCGATGTCGATTCGTGAGCTGAGTGG 555
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 603 TyrAspAspGlyLysThrTyrHisVal-----GlyGluGlnTrp 615
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 14
US-10-995-561-625
; Sequence 625, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 625
; LENGTH: 1259
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-625

Alignment Scores:
Pred. No.: 1.56e-05 Length: 1259
Score: 166.50 Matches: 36
Percent Similarity: 44.83% Conservative: 16
Best Local Similarity: 31.03% Mismatches: 41
Query Match: 5.31% Indels: 24
Gaps: 4

US-09-940-235-11 (1-1782) x US-10-995-561-625 (1-1259)
QY 254 CCCATAGCTGAGAAAGTTTGTATCATGCTGAGCACTTCTATGTGTGAGAAACG 313
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1074 ProthraspaserCyspHeasprProTyThrValSerHisTyrAlaValGlyAspGlu 1093
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 314 TGGAGAA-----GGCAGC 327
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1094 TrpGluArgMetSerGluSerGlyPheLeuLeuGlyGlnCysGlnCysLeuGlyPheGlySer 1113
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 328 GAGCGCATCTGCACTTCTAGAAATATGATGCATGACATGACAGACACATCCTAT 387
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1114 GlyHisPheArgCysAspSerSerArgTyrCysHisAspAsnGly-----ValAsnTyr 1131
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 388 AGAATTGAGACACCTGAGACAGAAAGATATATGAGAAACCTGCTCCAGTCATCTGC 447
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1132 LysIleGlyGlyLysTrpAspArgGlnGlyGluAsnGlyGlnMetMetSerCysThrCys 1151
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

QY 448 ACAAGCAACGCCGAGAGAGATGGAAGTGTGAGAGGACACCTCTGTGACACCAATCG 507
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1152 LeuGlyAsnGlyLysGlyGlnPheLysCysAspProHis-----GluAlaThrCys 1168
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 508 AGCGGATCTGGCCCTTCAACCGATGTCGATTCGTGAGCTGAGTGG 555
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1169 TyrAspAspGlyLysThrTyrHisVal-----GlyGluGlnTrp 1181
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 15
US-10-995-561-630
; Sequence 630, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 630
; LENGTH: 1315
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-630

Alignment Scores:
Pred. No.: 1.57e-05 Length: 1315
Score: 166.50 Matches: 36
Percent Similarity: 44.83% Conservative: 16
Best Local Similarity: 31.03% Mismatches: 41
Query Match: 5.31% Indels: 24
Gaps: 4

US-09-940-235-11 (1-1782) x US-10-995-561-630 (1-1315)
QY 254 CCCATAGCTGAGAAAGTTTGTATCATGCTGAGCACTTCTATGTGTGAGAAACG 313
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1130 ProthraspaserCyspHeasprProTyThrValSerHisTyrAlaValGlyAspGlu 1149
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 314 TGGAGAA-----GGCAGC 327
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1150 TrpGluArgMetSerGluSerGlyPheLeuLeuGlyGlnCysGlnCysLeuGlyPheGlySer 1169
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 328 GAGCGCATCTGCACTTCTAGAAATATGATGCATGACATGACAGACACATCCTAT 387
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1170 GlyHisPheArgCysAspSerSerArgTyrCysHisAspAsnGly-----ValAsnTyr 1187
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 388 AGAATTGAGACACCTGAGACAGAAAGATATATGAGAAACCTGCTCCAGTCATCTGC 447
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1188 LysIleGlyGlyLysTrpAspArgGlnGlyGluAsnGlyGlnMetMetSerCysThrCys 1207
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 448 ACAAGCAACGCCGAGAGAGATGGAAGTGTGAGAGGACACCTCTGTGACACCAATCG 507
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1208 LeuGlyAsnGlyLysGlyGlnPheLysCysAspProHis-----GluAlaThrCys 1224
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 508 AGCGGATCTGGCCCTTCAACCGATGTCGATTCGTGAGCTGAGTGG 555
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1225 TyrAspAspGlyLysThrTyrHisVal-----GlyGluGlnTrp 1237
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Search completed: January 28, 2006, 02:58:11
 Job time : 24.1597 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 28, 2006, 01:49:35 ; Search time 27.3089 Seconds

(without alignments)
12556.953 Million cell updates/sec

Title: US-09-940-235-11

Perfect score: 3134
Sequence: 1 tcgtctcacgtctcgtcgcg.....ataacctaacgacaataa 1782

Scoring table: BLOSUM62

Xgapop 10.0	Xgapext 0.5
Ygapop 10.0	Ygapext 0.5
Fgapop 6.0	Fgapext 7.0
Delop 6.0	Delext 7.0

Searched: 283416 seqs, 9621673 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

```
-MODEL=frame+ n2p.model -DBV=xld  
-O=/cgn2.1/USPNO_spool_P/US09940235/runat_27012006_144218_27578/app_query.fasta_1.7708  
-DB=PIR -QFMT=fastan -SUFRTX=tpir -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=bl0sum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR SCORE=pcr -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pcr -NORM=ext -HEASize=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09940235 @CCN 1.1 185 @runat_27012006_144218_27578 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEDEBERT -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database :

```
PIR 80:**  
1: pir1:**  
2: pir2:**  
3: pir3:**  
4: pir4:**
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2158	68.9	440	1 A22801	streptokinase prec
2	2120.5	67.7	415	1 BZSO	streptokinase A (B)
3	2116	67.5	440	2 S02723	streptokinase G pr
4	1938	61.8	440	2 S02724	streptokinase A pr
5	1841	58.7	414	2 JU0292	streptokinase - St
6	1838	58.6	440	2 S04168	streptokinase A pr
7	788.5	25.2	197	2 S53334	streptokinase - St
8	620	19.8	128	2 S77671	streptokinase A (E)
9	569	18.2	128	2 S77680	streptokinase A (E)
10	557	17.8	128	2 S77688	streptokinase A (B)
11	552	17.6	128	2 S77679	streptokinase A (B)
12	503	16.0	2386	1 FNHU	fibronectin precur
13	481	15.3	2265	1 FNBO	fibronectin - bov1
14	454.5	14.5	2477	2 S14428	fibronectin precur

15	423	13.5	128	2 S77676	streptokinase A (B)
16	420.5	13.4	2481	2 A43908	fibronectin - Afr1
17	419	13.4	128	2 S77673	streptokinase A (B)
18	419	13.4	128	2 S77687	streptokinase A (B)
19	417	13.3	128	2 S77682	streptokinase A (B)
20	416	13.3	128	2 S77678	streptokinase A (B)
21	408	13.0	128	2 S77677	streptokinase A (B)
22	407	13.0	128	2 S77685	streptokinase A (B)
23	401	12.8	128	2 S77683	streptokinase A (B)
24	398	12.7	128	2 S77686	streptokinase A (B)
25	397	12.7	128	2 S77684	streptokinase A (B)
26	396	12.6	128	2 S77681	streptokinase A (B)
27	390	12.4	128	2 S77674	streptokinase A (B)
28	386	12.3	128	2 S77675	streptokinase A (B)
29	147	4.7	1020	2 A29355	fibronectin - chic
30	140.5	4.5	2231	2 D71870	hypothetical prote
31	136.5	4.4	13055	2 T16580	hypothetical prote
32	127.5	4.1	928	2 C81265	probable lipoprote
33	127.5	4.1	978	2 D81411	probable lipoprote
34	127	4.1	1965	2 S75200	fat protein - Syne
35	127	4.1	2256	2 A01018	large repetitive p
36	126	4.0	1019	2 T50251	hypothetical coile
37	125	4.0	1417	2 H90670	probable adhesin l
38	125	4.0	1417	2 D85521	probable adhesin e
39	125	4.0	2140	2 P95074	serine proteinase,
40	124.5	4.0	190	2 I51279	fibronectin - east
41	124.5	4.0	617	2 D66978	hypothetical prote
42	124	4.0	1002	2 T02458	hypothetical prote
43	124	4.0	3394	2 T18501	hypothetical prote
44	123.5	3.9	1176	2 T47444	hypothetical prote
45	123	3.9	1286	2 T16507	hypothetical prote

ALIGNMENTS

RESULT 1

A22801

streptokinase precursor - Streptococcus "equisimilis"

C/Speclab: Streptococcus "equisimilis"

C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004

C/Accession: A22801

R/Malke, H.; Roe, B.; Ferretti, J.J.

Gene 34, 357-362, 1985

A/Title: Nucleotide sequence of the streptokinase gene from Streptococcus equisimilis H

A/Reference number: A22801, MUID:85232082; PMID:2989113

A/Accession: A22801

A/Molecule type: DNA

A/Residues: 1-440 <MAL>

A/Cross-references: UNIPROT:P00779; UNIPARC:UPI000002BE73; GB:X72832; NID:9407876; PIDN

A/Experimental source: strain H46A

A/Genetics:

C/Superfamily: streptokinase

Alignment Scores:

Pred. No.:	6.23e-140	Length:	440
Score:	2158.00	Matches:	414
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	68.86%	Indels:	0
DB:	1	Gaps:	0

US-09-940-235-11 (1-1782) x A22801 (1-440)

QY	538	ATTGCTGACGTGAGTGGCTGTCGACGTCATCTGTCAACACAGCAATGGTTGTT	597
DB	27	IIehlaelyProglutritrpleuleuAspArgProSerValbnaSnSerGlnleValVal	46
QY	598	AGCGTTCCTGTACTGTGAGGGAGCAATCAAGACATTAGTTAAATTTTGAATC	657
DB	47	SerValaIahelyhrValGluGlyThrAsnGlnAspIleSerleuysrPhePheGluIle	66
QY	658	GATCTAACATCAGCACTGCTCATGTGAGGAAAGACAGACAAAGGCTTAAGTCCAAAATCA	717

```

Db      67  AspleuThrSerArgProIahIsgIyLysThrGluGlnGlyLeuSerProlySser 86
Qy      718  AAACCATTTGCTAGTATGTTGGCGCGGATGTGCATATACTTGAGAAAGCTGACTTACTA 777
Db      87  LysProPheAlaThrAspSerGlyAlaMetSerHisIleuGlnIulysAlaAspLeuLeu 106
Qy      778  AAGGATATTCAGAACAAATGATTCGTACGTCCAGTAAAGAGACGACACTTGTGAGTC 837
Db      107  LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 126
Qy      838  ATTGATTTTGCAGAGATGCAACCACTTACTGATTCGAACGAGCAAGCTCTTACTTGCTGAC 897
Db      127  IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValIlyrPheAlaAsp 146
Qy      898  AAAAGATGGTTCGGTAACTTCCGCGACCAACTGTCTCAAGAAATTTTGTACCGGACAT 957
Db      147  LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 166
Qy      958  GTGCGCCGTTAGCCATATAAAGAAAAACCAATTCAAAAACCAAGCCGAAATCTGTGATGTG 1017
Db      167  ValArgValAlaArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 186
Qy      1018  GAATATACCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTGACAGGCTCAAA 1077
Db      187  GluTyrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuLys 206
Qy      1078  GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATGACATCTCAAGAAATTTACTA 1137
Db      207  AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 226
Qy      1138  GCTCAGACGCAAAAGATTTTAAACAAAAACCAACCGCTATTCAGATTTTATGAACTGAC 1197
Db      227  AlaGlnAlaGlnSerIleLeuAsnLysAsnHisAspGlyTyrThrIleTyrGluAspAsp 246
Qy      1198  TCCCTCAATCGTCACCTCATGACATGATCTTCCGACATTTTCCGATTTTAACTGATCAAGAG 1257
Db      247  SerSerIleValThrHisAspAsnAspIlePheAspGlnThrIleLeuProAspGlnGln 266
Qy      1258  TTTACTTACCGTGTAAAAAATCGGGAACAAGCTTATAGGATCAATAAAAATCTGGTCTG 1317
Db      267  PheThrTyrArgValLysAsnAspGlnGlnAlaTyrArgIleAsnLysLysSerGlyLeu 286
Qy      1318  AATGAGAAATTAACAACAACCTGATCTGTGAGAAATTTATGCTCTTAAAAAAGG 1377
Db      287  AsnGluGlnIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 306
Qy      1378  GAAAGCGGTAATGATCCCTTGAATGCGACATCTGGAATCTGTCCATCAAAATACGTT 1437
Db      307  GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326
Qy      1438  GATGTGATACCAAGAAATTTGCTAAAAAGTGAGACGCTTTAAACAGCTACCGAAGCTAAC 1497
Db      327  AspValAspThrAsnGlnLeuLeuLysSerGlnGlnLeuLeuThrAlaSerGlnAlaGln 346
Qy      1498  TTAAGCTTCAAGATTTTATGATCAATCTCTGTGATTAAGGCTTAACTACTTCAACAATTC 1557
Db      347  LeuAspPheArgAspLeuLysAspProArgAspLysAlaLysLeuLeuLysAsnLeu 366
Qy      1558  GATGCTTTTGGTATTAAGCATATACCTTAACTGGAAGAAAGTAGAGATATACAGATGAC 1617
Db      367  AspAlaPheGlyLysIleMetAspTyrThrLeuThrGlyLysValGlnAspAsnHisAspAsp 386
Qy      1618  ACCAACCTATCATTAACCGTTATATATGGCAAGCGACCGAAGAGAGAGATGCTATCAT 1677
Db      387  ThrAsnAlaGlnIleThrValIlyrMetGlyLysArgProGlnGlyGlnAsnAlaSerTyr 406
Qy      1678  CATTAGCTTATGATTAAGATCGTTATATCCGAAGAAAGCAAGAACTTTTACAGCTTACG 1737
Db      407  HisLeuAlaTyrAspLysAspArgTyrThrGlnGlnGlnLysArgLysValIlyrSerTyrLeu 426
Qy      1738  CGTTATACAGGACACCTTACTGATTAACCTTAAGACAA 1779

```

```

Db      427  ArgTyrThrGlyThrProIleProAspAsnProAsnAspLys 440
RESULT 2
BZSO
Streptokinase (EC 3.4.-.-) - Streptococcus sp.
C/Species: Streptococcus sp.
C/Date: 05-Apr-1983 #sequence_revision 05-Apr-1983 #text_change 09-Jul-2004
C/Accession: A00967
R/Jackson, K.W.; Tang, J.
Biochemistry 21, 6620-6625, 1982
A/Title: Complete amino acid sequence of streptokinase and its homology with serine protease
A/Reference number: A00967; PMID:83127125; PMID:6760891
A/Accession: A00967
A/Molecule type: protein
A/Residues: 1-415 <0AC>
A/Cross-references: UNIPROT:P00779; UNIPARC:UPI0000172BF5
A/Note: 169-Asp and 181-Asp were also found
A/Note: this protein is not a protease, but it activates plasminogen by complexing with C/superfamily: streptokinase
C/Keywords: hydrolase

Alignment Scores:
Pred. No.: 2,27e-137 Length: 415
Score: 2120.50 Matches: 410
Percent Similarity: 99.28% Conservative: 2
Best Local Similarity: 98.80% Mismatches: 2
Query Match: 67.66% Indels: 1
DB: 1 Gaps: 1

US-09-940-235-11 (1-1782) x BZSO (1-415)

Qy      538  ATTGCTGACCTGAGTGGCTGTGACCCGTCATCTGTCAACAACAGCAATGGTTGTT 597
Db      1  IleAlaGlyProGluThrLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
Qy      598  AGCGTTCGCTGCTGTGTTGAGGGGAGCAATCAAGACATTAAGTCTTAAATTTTGAATATC 657
Db      21  SerValAlaGlyThrValGlnGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 40
Qy      658  GATCTAATCATCAAGACCTGTCTCATGAGAGAAAGACAGAGCAAGGCTTAACTCCAAATCA 717
Db      41  AspleuThrSerArgProIahIsgIyLysThrGluGlnGlyLeuSerProlySser 60
Qy      718  AAACATTTGCTACTGATATGTTGGCGGATGTGCATATACTTGAGAAAGCTGACTTACTA 777
Db      61  LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlnLysAlaAspLeuLeu 80
Qy      778  AAGGCTATTCAGAACAAATTTGATCGCTAAGTCCACAGTAAAGACGACTTGTGAGGTC 837
Db      81  LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100
Qy      838  ATTGATTTTGCAGAGATGCAACCACTTACTGATTCGAACGAGCAAGCTCTTACTTGCTGAC 897
Db      101  IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValIlyrPheAlaAsp 120
Qy      898  AAAAGATGGTTCGGTAACTTCCGCGACCAACTGTCTCAAGAAATTTTGTGAGCGGACAT 957
Db      121  LysAspGlySerValThrLeuProThrGlnProValGlnGlnPheLeuLeuSerGlyHis 140
Qy      958  GTGCGCGTTAGACCATATAAAGAAAAACCAATTCAAAAACCAAGCAAAATCTGTGAGAGTG 1017
Db      141  ValArgValAlaArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
Qy      1018  GAATATACCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTGACAGGCTCAAA 1077
Db      161  GluTyrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuLys 180
Qy      1078  GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATGACATCTCAAGAAATTTACTA 1137
Db      181  LeuThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeuLeu 200
Qy      1138  GCTCAGACGCAAAAGATTTTAAACAAAAACCAACCGCTATTCAGATTTTATGAACTGAC 1197

```


Db 201 AAGlnAlaGlnSerIleuLeuAsnIleProGlyThrIleTyrgluAsp 220
Qy 1198 TCCTCAATGTCATCATGATGACATATTTCCGTAAGATTTTACCATGATGATCAAG 1257
Db 221 SerSerIleValIleThrIleAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
Qy 1258 TTTTACTTACCGGTGTTAAAAATCGGGAACAAGCTTTATAGATCAATAAAAAATCGTCTG 1317
Db 241 PheThrTyArgValIleAsnAsnArgGlnGlnAlaTyArgIleAsnIleAsnIleuSylSerGlyLeu 260
Qy 1318 AATGAAGAATAAACAACATGACCTGATCTCGAGAAATAT--TACGTCCTTAAAAA 1374
Db 261 AangIuglnIleAsnAsnThrAspIleSerLeuGlnIlyrIlySerTyValIleuSyls 280
Qy 1375 GGGGAAAAGCCGATGATCCCTTGTGATGCGATCACTGAAACGTTCCACATCAATAC 1434
Db 281 GlyIlyuysProTyArgProPheAspArgSerHisIleuSylAsnPheThrIleuSylr 300
Qy 1435 GTTGATGTGATACCAACGAATTTGTAATAAGTAGACAGCTTTAACAGCTAGGAAAGCT 1494
Db 301 ValAspValAspThrAsnGlnIleuLeuSylSerGlnIleuLeuThrIleAspGlnArg 320
Qy 1495 AACTTAGACTTCAGAGATTTATTCGATCTCGTATAGAGCTTAACTACTCTACACAT 1554
Db 321 AsnIleAspPheArgAspIleuTyArgProArgAspIlysalIlySylleuIlyTyAsnAsn 340
Qy 1555 CTGCAATGCTTTGGTATTATGACCTTATACCTTAACTGGAAGAAGTAGAGATTCACGAT 1614
Db 341 LeuAspAlaPheGlyIleMetAspTyThrIleuThrGlyIlyValGluAspAsnHisAsp 360
Qy 1615 GACACCAACCGTATCATTAACCGTTTATATGAGCAAGCAAGCAAGAGAGAAATGCTAGC 1674
Db 361 AspThrAsnAspGlyIleIleThrValTyMetGlyIlyAspProGlnGlnIlyuAsnAlaSer 380
Qy 1675 TATCATTTAGCTATGATTAAGAATGCTTATACCGAAAGAAGCAAGAGTTTACAGTAC 1734
Db 381 TyrHisIleuAlaIlyTyArgAspAspArgTyThrGlnGlnIlyuArgIlyValTySerTy 400
Qy 1735 CTGGCTTATACAGGAGCACCTTATACCTGATTAACCTTAACGACAAA 1779
Db 401 LeuArgTyThrGlyThrProIleProAspAsnProAspAspIlyS 415
RESULT 3
S02723
Streptokinase G precursor - Streptococcus sp. (strain 19908)
C:Species: Streptococcus sp.
C:Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 09-Jul-2004
C:Accession: S02723
R:Walter, F.; Siegel, M.; Malke, H.
Nucleic Acids Res. 17, 1262, 1989
A:Title: Nucleotide sequence of the streptokinase gene from a group-G Streptococcus.
A:Reference number: S02723; MID:89160265; PMID:2922269
A:Accession: S02723
A:Molecule type: DNA
A:Residues: 1-440 <MAL>
A:Cross-references: UNIPROT:P10519; UNIPARC:UPI000013614B; EMBL:X13400; NID:947095; PIDD
C:Genetics:
A:Gene: skg
C:Superfamily: streptokinase
F:1-26/Domain: signal sequence #status predicted <Sig>
F:27-440/Product: streptokinase #status predicted <Mat>
Alignment Scores:
Pred. No.: 4,64e-137 Length: 440
Score: 2116.00 Matches: 406
Percent Similarity: 98.55% Conservative: 2
Best Local Similarity: 98.07% Mismatches: 6
Query Match: 67.52% Indels: 0
DB: 2 Gaps: 0
US-09-940-235-11 (1-1782) x S02723 (1-440)
Qy 538 ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCAATTGTTGTT 597

Db 27 IleAlaGlyProGlnIlyrIleuLeuAspArgProSerValAsnAsnSerGlnIleValAl 46
Qy 598 AGCGTCCTGCTACTGTGGAGGGAGCAATGAAACATTAATCTTAATTTTGAATC 657
Db 47 SerValAlaGlyThrValGlnGlyThrAsnIleAspIleSerLeuSylPhePheGlnIle 66
Qy 658 GATCAATCAATCAAGCACTGCTCATGAGAGAAAGACAGAGCAAGGCTTAAGTCAAAATCA 717
Db 67 AspIleuThrSerArgProAlaHisGlyGlyIlySerThrGlnGlnIlyuSerProIlySer 86
Qy 718 AAACCATTTGCTACTGATAGTGCGCGATGTCACATAACTTGAGAAAGCTGACTTACTA 777
Db 87 LysIleuPheAlaThrAspSerGlyAlaMetProHisIlySylleuGlnIlyValAspIleu 106
Qy 778 AAGCTTATCAAGAAACAATGATGCTTAAGCTCAACGATTAACAGCACTTGTGAGTC 837
Db 107 LysAlaIleGlnGlnIlyuIleAlaAsnValHisSerAsnAspAspTyPheGlnVal 126
Qy 838 ATTGATTTTGGCAAGTGAACCATTAATGATGAAGCAAGGCAAGTCTACTTGTGTCAG 897
Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyIlyValTyPheAlaAsp 146
Qy 898 AAAGATGTTTCGATTAACCTTGCCGACCAACCTGTCCAAGATTTTGTCTAAGCGACAT 957
Db 147 LysAspGlySerValThrIleuProIleGlnProValGlnIlyuPheIleuIlyGlyHis 166
Qy 958 GTGCGCGTTAGACCATTAATAAGAAAAACAATACAAACAAGCAAGAAATCTGTTGATGTG 1017
Db 167 ValArgValArgProTyIlySylIlyAspProValGlnAsnGlnAlaIlySerValAspVal 186
Qy 1018 GAATATATCTGATGACGTTTACCTCCCTTAAACCTGATGACGATTTACAGCCAGGCTCAAA 1077
Db 187 GluTyThrValGlnPheThrProLeuAsnProAspAspPheArgProAlaIleuSyls 206
Qy 1078 GATTAAGGTAATTTGAAAAACATGACTGTCGTGACACATCAATCATCTCAAGATTTACTA 1137
Db 207 AspThrIlySylleuSylThrIleuAlaIleGlyAspThrIleThrSerGlnIlyleu 226
Qy 1138 GCTCAAGCAACAAGCATTTTAAACAACCAACCAAGGCTTATAGCATTTATGAACGTGAC 1197
Db 227 AlaGlnAlaGlnSerIleLeuAsnIlyAsnHisProGlyTyThrIleTyrgluAsp 246
Qy 1198 TCCTCAATGCTCATGATGACATGACATTTTCCGTAAGATTTTACCAATGATGACAGAG 1257
Db 247 SerSerIleValIleThrIleAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
Qy 1258 TTTTACTTACCGGTGTTAAAAATCGGGAACAAGCTTATAGATCAATAAAAAATCGTCTG 1317
Db 267 PheThrTyrHisValIlyAsnAsnArgGlnGlnAlaTyArgIleAsnIlySylSerGlyLeu 286
Qy 1318 AATGAAGAATAAACAACATGACCTGATCTCGAGAAATATTAACGTCCTTAAAAAAGG 1377
Db 287 AsnGlnGlnIlyleAsnAsnThrAspIleIleSerIlyuSylTyrgValIleuSylGly 306
Qy 1378 GAAAAAGCGATGATCCCTTGTGATCGAGTCACCTTGAACCTGTCACCATCAATACGTT 1437
Db 307 GluIlySylProTyArgProPheAspArgSerHisIleuSylAsnPheThrIleuSylVal 326
Qy 1438 GATGCTATGCAACGAATTTGTAATAAGTAGACAGCTTTAACAGCTAGGAAAGCTTACTA 1497
Db 327 AspValAsnThrAsnGlnIlyleuSylSerGlnIlyleuIlyThrIleAspGlnArgAsn 346
Qy 1498 TTAGACTTCAGAGATTTATGATCCCTGATAGGCTTAAAGCTTACATCAACATCTC 1557
Db 347 LeuAspPheArgAspIleuTyArgProArgAspIlysalIlySylleuIlyTyArgAsnIleu 366
Qy 1558 GATGCTTTGGTATTATGACCTATACCTTAACTGGAAGAAGTAGAGATTAATCAGATGAC 1617
Db 367 AspAlaPheGlyIleMetAspTyThrIleuThrGlyIlyValGluAspAsnHisAspAsp 386
Qy 1618 ACCAACCGTATCAATTAACCGTTTATATGAGCAAGGACCCGAAGAGAGATTCGACTAT 1677

Db 387 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGluAsnAlaSerTyr 406
QY 1678 CATTTAGCCTATGATTAAGATGCTTATACCGAAGAAAGAGACTTTACACTGCTG 1137
Db 407 HisLeuAlaTyrAspLysAspArgTyrThrGluGluArgGluValTyrSerTyrLeu 426
QY 1738 CGTTATACAGGAGCACCTATACCTGATTAACCTTAACGACAAA 1779
Db 427 ArgTyrThrGlyThrProIleProAspAsnProAsnAspLys 440
RESULT 4
S02724
streptokinase A precursor - Streptococcus pyogenes (strain SF130/13)
C:Species: Streptococcus pyogenes
C>Date: 18-Oct-1989 #sequence_revision 18-Oct-1989 #text_change 09-Jul-2004
C:Accession: S02724
R:Walter, F.; Siegel, M.; Malke, H.
Nucleic Acids Res. 17, 1261, 1989
A:Title: Nucleotide sequence of the streptokinase gene from a Streptococcus pyogenes lys
A:Reference number: S02724; PMID:89160264; PMID:2646590
A:Accession: S02724
A:Molecule type: DNA
A:Residues: 1-440 <MAL>
C:Genetics:
A:Gene: ska
C:Superfamily: streptokinase
F:1-26/Domain: signal sequence #status predicted <Sig>
F:27-440/Product: streptokinase #status predicted <MAT>
Alignment Scores:
Pred. No.: 6 94e-125 Length: 440
Score: 1938.00 Matches: 366
Percent Similarity: 93.72% Conservative: 22
Best Local Similarity: 88.41% Mismatches: 26
Query Match: 61.84% Indels: 0
DB: 2 Gaps: 0
US-09-940-235-11 (1-1782) x S02724 (1-440)
QY 538 ATTGCTGACCTGATGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCAATGTTGTT 597
Db 27 IleAlaGlyTyrGlyTyrIleuProAspArgProProIleAsnAsnSerGlnLeuValAla 46
QY 598 AGCGTGTCTGACTGTTGAGGAGGAGCAATCAAGACATTAAGTCTTAATTTTGAATC 657
Db 47 SerMetAlaGlyIleValGluGlyThrAspLysValPheIleAsnPheGluIle 66
QY 658 GATCTAACATCAAGACCTGCTCATGAGGAGAAAGACAGCAAGGCTTAATGCCAATATCA 717
Db 67 AspIleuThrSerGlnProIleAsnGlyLysThrGluGlnGlyLeuSerProLysSer 86
QY 718 AAACCATTTGCTACTGATGATGAGCGGATGTCACATTAATCTTGAGAAAGCTGACTTACTA 777
Db 87 LysProPheAlaThrAspAsnGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu 106
QY 778 AAGGCTATTCAAGAACATTTGATCGCTTACGTCACAGTACGACGACTTACTTGAAGTC 837
Db 107 LysAlaIleGlnIleuLeuIleAlaAsnValHisSerAsnAspGlyTyrPheAlaAsp 126
QY 838 ATTGATTTTGGCAAGATGCAACCATTAATCTGATCGAAAGGCAAGGCTTACTTGTGATC 897
Db 127 IleAspPheAlaSerAspAlaThrIleThrAspAsnGlyLysValTyrPheAlaAsp 146
QY 898 AAAGATGCTTGGTAACTTGGCGACCGAACCTGTCACAGATTTTGGTGAAGCGGACAT 957
Db 147 LysAspGlySerValThrIleuProThrGlnProValGlnGluPheLeuValLysGlyHis 166
QY 958 GTGCGCGTTAGACCATATTAAGAAAAACCAATACAAACCAAGCGAATCTGTTGATGTC 1017
Db 167 ValArgValArgProTyrLysGlyLysProValGlnAsnGlnAlaLysSerValAspVal 186
QY 1018 GAATATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTCAGACCAAGGCTTCMAA 1077

Db 187 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheAspProGlyLeuLys 206
QY 1078 GATACTTAAGCTATTGAAACACTAGCTATGCTGACCAACATCATCTCAAGAAATTACTA 1137
Db 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnIleLeu 226
QY 1138 GCTCAAGCAAGACATTTTAAACAAACACCCAGGCTATCGATTTTGAACGAC 1197
Db 227 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp 246
QY 1198 TCCTCAATCGTCACTCATGACCAATGACATTTTCGTCAGATTTTACCAAGATCAAGAG 1257
Db 247 SerSerIleValIleThrHisAspAsnAspIlePheThrIleLeuProMetAspGln 266
QY 1258 TTTACTTACCGGTGTTAAAAATCGGAAACAAGCTTATAGATCAATAAAAATCTGCTG 1317
Db 267 PheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsnProLysThrGlyIle 286
QY 1318 AATGAAGAAATTAACCAACTGACCTGATCTCTGAGAAATTTACGTCCTTAAAAAGG 1377
Db 287 LysGluLysThrAsnAsnThrAspLeuValSerGluLysTyrValLeuLysGlnGly 306
QY 1378 GAAAGCCGTATGATCCCTTGAATCGGACGACCTTGAACCTGTTCAACATCAATACGTT 1437
Db 307 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326
QY 1438 GATGTCGATACCAACGAATGCTTAAAGTAGACGCTTAAACGCTAGCAAGCTAAC 1497
Db 327 AspValAsnThrAsnGlnLeuLeuLysSerGlnGlnLeuThrAlaSerGluGlyAsn 346
QY 1498 TTAGACTTCAGAGATTTATTCGATCCTCGTATAGGCTTAACTACTTCAACAATCTC 1557
Db 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLysTyrAsnAsnLeu 366
QY 1558 GATGCTTTGGTATTTATGACCTTACCTTAACTGGAAGATGAGATTAATCAAGATAC 1617
Db 367 AspAlaPheAspIleMetAsnTyrThrLeuThrGlyLysValGluAspAsnHisAspLys 386
QY 1618 ACCAACCGTATCAATTAACCGTTTATATGGGCAAGCAACCGAAGAGAAATGCTAGCTAT 1677
Db 387 AsnAsnArgValAlaThrValTyrMetGlyLysArgProLysGlyAlaLysGlySerTyr 406
QY 1678 CATTTAGCCTATGATTAAGATGCTTATACCGAAGAAAGACGAGAAGTTTACAGTACTG 1737
Db 407 HisLeuAlaTyrAspLysAspLeuTyrThrGluGluGluArgLysAlaTyrSerTyrLeu 426
QY 1738 CGTTATACAGGAGCACCTATACCTGATTAACCTTAACGACAAA 1779
Db 427 ArgTyrThrGlyThrProIleProAspAsnProAsnAspLys 440
RESULT 5
J00292
streptokinase - Streptococcus pyogenes
C:Species: Streptococcus pyogenes
C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C:Accession: A43867; J00292
R:Okuni, H.; Todome, Y.; Suzuki, H.; Mizuse, M.; Kotani, N.; Horiiuchi, K.; Shikama, N.;
Infect. Immun. 60, 278-283, 1992
A:Title: Immunochemical studies and complete amino acid sequence of the streptokinase fr
A:Reference number: A43867; PMID:92104686; PMID:11370275
A:Accession: A43867
A:Molecule type: protein
A:Residues: 1-414 <OHK>
A:Cross-references: UNIPROT:Q57391; UNIPARC:UPI0000175C82
A:Experimental source: M type 12 strain A374
A:Note: sequence extracted from NCBI backbone (NCBIP:74592)
C:Superfamily: streptokinase
Alignment Scores:
Pred. No.: 2 97e-118 Length: 414
Score: 1841.00 Matches: 354
Percent Similarity: 91.06% Conservative: 23

Best Local Similarity: 85.51% Mismatches: 37
 Query Match: 58.74% Indels: 0
 DB: 2 Gaps: 0

US-09-940-235-11 (1-1782) x JU0292 (1-414)

```

QY 538 ATTTCGACCTGAGCGCTGCTAGACCGTTCATCTGTCAACAACAGCCAAATGGTTGT 597
DB 1 TlealaglyProdluttrleuAspArgProSerValAsnAsnSerGlnLeuVal 20
QY 598 AGCGTTCGTGTACTGTGAGGGGAGCAATCAAGACATTAAGCTTAATTTTGAATC 657
DB 21 SerValalaglyThrValGlnGlyThrAsnGlnGlnLeuSerLeuysPhepGlnIle 40
QY 658 GATTCATACATCAAGACCTGCTCATGAGAGAAAGACAGACAGGCTTAAGTCAAAATCA 717
DB 41 AspLeuThrSerArgProAlaHisGlyGlyLeuThrGlnGlnGlnLeuSerProLysSer 60
QY 718 AAACCATTTGCTACTGATAGTGGCGGAGTGCATTAACCTTGAGAAAGCTGACTTA 777
DB 61 LysProPheAlaThrAsnSerSerAlaMetProHisLysLeuGlnLysValAspLeuLeu 80
QY 778 AAGCGTATTCAGAAACAATTAATGCTGCTAAGTCAACGATCAAGCACTATTGAGTGC 837
DB 81 LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGlnVal 100
QY 838 ATTGATTTTGAAGCGATGCAACCATTAATCATGATGCAAGCGCAAGGTCTATTGCTGAC 897
DB 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
QY 898 AAAGATGATTTCGGTAACTTTCGCCAGCCCAACTGTCACAGAAATTTTGTAAAGCGACAT 957
DB 121 ArgAspAspSerValThrLeuProThrGlnProValGlnGlnLeuLeuSerGlyHis 140
QY 958 GTGCGCGTTAGACCATATTAAGAAAAACCAATACAAACCAAGCAAAATGTTGATGTCG 1017
DB 141 ValArgValArgProTyrGlnProLysValAlaValHisAsnSerAlaGlnArgValAsnVal 160
QY 1018 GAATTAAGTACAGTTCAGTTTACTCCCTTAACCCGATGACGATTTTCAAGACGCTCAAA 1077
DB 161 AsnTyrGlnValSerPheValSerGlnThrGlyAsnLeuAspPheThrProSerLeuLys 180
QY 1078 GATTAAGTACGTTTGAACCACTAGCTATCGGTGACACACATCATCTCAAGATTACTA 1137
DB 181 GlnArgTyrHisIleuThrThrIleuAlaValGlyAspSerLeuSerGlnGlnLeuAla 200
QY 1138 GCTCAAGCAACAAGCATTTTAAACAACCAACCCAGGCTATACGATTTTGAACGTCAC 1197
DB 201 AlaIleAlaGlnPheIleLeuSerLysGlnHisProAspTyrIleIleThrLysAsp 220
QY 1198 TCCCTCAATGTCATCATGACATGACATTTTCCGTAACGATTTTACCATGATCAAGAG 1257
DB 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGln 240
QY 1258 TTACTTACCGTGTAAATAATGGGAAACAAGCTTATAGATCAATTAATAAATCGTGGC 1317
DB 241 PheThrTyrHisIleLysAspArgGlnGlnAlaTyrLysAlaAsnSerLysThrGlyIle 260
QY 1318 AATGAAGAATTAACAACAACATGACCTGATCTGTAGAATATTAAGTCTTAATAAAAGG 1377
DB 261 ValGlnLysThrAsnAsnThrAspLeuIleSerGlnLysTyrTyrValLeuLysGly 280
QY 1378 GAAAAGCGGTATGATCCCTTTGATGCGAGTCACTTGAATCTGTCAACATCAATAGCTT 1437
DB 281 GlnGlnProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
QY 1438 GATTCGATTAACAAGAAATTTGCTAAATAATGAGCACTCTTAACAGCTAGCAAGCTAAC 1497
DB 301 AspValAspThrAsnGlnLeuLeuLysSerGlnGlnLeuThrAlaSerGlnLysAsn 320
QY 1498 TTAGACTTCAGAGATTTTATACGATCTCGGATTAAGGCTTAACCTATCTTCAACAATTC 1557
DB 321 LeuAspPheAspArgAspLeuTyrAspProAlaGlyAspLysValAlaLysLeuLeuLys 340
  
```

```

QY 1558 GATGCTTTGGTATTAATGACCTTAATACCTTAATGGAAGTAGAGATTAATCAGATGAC 1617
DB 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGlnAspAsnHisAsnAsp 360
QY 1618 ACCAAGCGTATCATTAACCGTTTATATGGGCAAGCCGCAAGAGAGAAATGCTAGTAT 1677
DB 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGlnGlnGlnValAlaSerTyr 380
QY 1678 CATTAAGCTTATGATTAAGATCGTTTATCCGAAGAAAGCAAGAAATTAAGCTTACCTG 1737
DB 381 HisLeuAlaTyrAspLysAspArgTyrThrGlnGlnGlnValArgGlnValTyrSerTyrLeu 400
QY 1738 CGTTATACAGGAGCAACCTATACCTGATTAACCTTAACGACAA 1779
DB 401 ArgTyrThrGlnTyrThrProIleProAspAsnProLysAspLys 414
  
```

RESULT 6

S04168
 streptokinase A precursor - Streptococcus pyogenes (strain NZ131)
 C:Species: Streptococcus pyogenes
 C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 09-Jul-2004
 C:Accession: S04168
 R:Huang, T.T.; Maik, H.; Ferretti, J.J.
 Mol. Microbiol. 3, 197-205, 1989
 A:Title: The streptokinase gene of group A streptococci: cloning, expression in Escheri-
 A:Reference number: S04168, MUID:89343623; PMID:2668686
 A:Accession: S04168
 A:Molecule type: DNA
 A:Residues: 1-440 <HUA>
 A:Cross-references: UNIPROT:Q57391, UNIPARC:UPI0000BD04A; EMBL:X51517; NID:g47437; PID
 A:Gene: ska
 C:Superfamily: streptokinase
 P:1-26/Domain: signal sequence #status predicted <SIG>
 F:27-440/Product: streptokinase A #status predicted <MAT>

Alignment Scores:

Pred. No.: 4.8e-118 Length: 440
 Score: 1838.00 Matches: 354
 Percent Similarity: 90.82% Conservative: 22
 Best Local Similarity: 85.51% Mismatches: 38
 Query Match: 58.65% Indels: 0
 DB: 2 Gaps: 0

US-09-940-235-11 (1-1782) x S04168 (1-440)

```

QY 538 ATTTCGACCTGAGCGCTGCTAGACCGTTCATCTGTCAACAACAGCCAAATGGTTGT 597
DB 27 TlealaglyTyrGlnuttrleuAspArgProSerValAsnAsnSerGlnLeuVal 46
QY 598 AGCGTTCGTGTACTGTGAGGGGAGCAATCAAGACATTAAGCTTAATTTTGAATC 657
DB 47 SerValalaglyThrValGlnGlyThrAsnGlnGlnLeuSerLeuysPhepGlnIle 66
QY 658 GATTCATACATCAAGACCTGCTCATGAGAGAAAGACAGACAGGCTTAAGTCAAAATCA 717
DB 67 AspLeuThrSerArgProAlaHisGlyGlyLeuThrGlnGlnGlnLeuSerProLysSer 86
QY 718 AAACCATTTGCTACTGATAGTGGCGGAGTGCATTAACCTTGAGAAAGCTGACTTACTA 777
DB 87 LysProPheAlaThrAspLysGlyAlaMetSerHisLysLeuGlnLysValAspLeuLeu 106
QY 778 AAGCGTATTCAGAAACAATTAATGCTGCTAAGTCAACGATCAAGCACTATTGAGTGC 837
DB 107 LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGlnVal 126
QY 838 ATTGATTTTGAAGCGATGCAACCATTAATGATGAAACGCAAGGTCTTATTGCTGAC 897
DB 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146
QY 898 AAAGATGATTTCGGTAACTTTCGCCAGCCCAACTGTCACAGAAATTTTGTAAAGCGACAT 957
  
```

Db	147	LysAspAspSerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis	166
Qy	958	GTGGCCGCTTGAACCACTATTAAAGAAAACCACATATACAAAACCAAGCGAAATCTGTTGATGTG	1017
Db	167	ValArgValLysProTyrGlnProLysIleValHisAsnSerIleGluArgValAsnIle	186
Qy	1018	GAATATACTGTACAGTTTACTCCCTTTAAACCCTGATGTCAGATTTCACAGACGGTCTCAAA	1077
Db	187	AsnTyrGluValSerPheValSerGluThrGlyAspLeuAspPheThrProLeuLeuArg	206
Qy	1078	GATATCAAGCTATTGAAAAACATAGCTATCGGTGACACCATCATCTCAAGAAATTACTA	1137
Db	207	AsnGlnTyrHisLeuThrThrLeuIleValGlyAspSerLeuSerSerGlnGluIleuVala	226
Qy	1138	GCTCAAGCAACAAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTATGAACGTAC	1197
Db	227	AlaIleAlaGlnPheIleLeuSerLysHisIleProAspTyrIleIleThrLysArgAsp	246
Qy	1198	TCCTCAATCGTCACTCATGACATGACATCATTTTCCGTACGATTTTACCAATGATCAAGAG	1257
Db	247	SerSerIleValThrHisAspAsnAspIlePheAsnThrIleLeuProMetAspGlnGlu	266
Qy	1258	TTTACTTACCGGTTTAAAAATCGGGAACACAGCTTATAGATCATAAAAATCTGGTCTG	1317
Db	267	PheThrTyrHisIleLysAspArgGlnGlnIleTyrLysAlaAsnSerLysThrGlyIle	286
Qy	1318	AATGAAGAAATTAACCAACATGACCTGTATCTCTGAGAAATTTACGCTTTAAAAAAGG	1377
Db	287	GluGluLysThrAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysIleGly	306
Qy	1378	GAAGAAGCCGTATGATCCCTTTGATGCGAGTCACTTGAACGTTCACCATCAAAATACGT	1437
Db	307	GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleAsnTyrVal	326
Qy	1438	GATGTGCAATCAACGAAATGCTGTAATAAGAGACGCTTTAACAAGCTAAGGGAACGTAC	1497
Db	327	AspValAsnThrAsnLysLeuLeuLysSerGlnGlnLeuLeuThrAlaSerGluArgAsn	346
Qy	1498	TTAAGCTTCAGAGATTTATACGATCCTCGTGATAGAGCTAAAGCTTACCAACATCTC	1557
Db	347	LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu	366
Qy	1558	GATGCTTTTGGTATTATGACTATACCTTAACTGGAAGAAAGTGAAGAAATATCAGATGAC	1617
Db	367	AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp	386
Qy	1618	ACCAACCGTATCATTAACCGTTTATATGCGGCAAGGCAACCGGAAGGAGAAAGCTAGCAT	1677
Db	387	ThrAsnArgIleIleThrValTyrMetGlyLysArgProGlnGlyGluAsnHisSerTyr	406
Qy	1678	CATTATAGCGTATGATAAAGATCGTTATACGAGAAGAAAGAGAGTTTACAGTACCTG	1737
Db	407	HisLeuAlaTyrAspLysAspArgTyrThrGlnGluGluArgGluValTyrSerTyrLeu	426
Qy	1738	CGTTATACAGGAGACCTATACCGTATACCTTAACCTTAACGACAA	1779
Db	427	ArgTyrThrGlyThrProIleProAspAsnProLysAspLys	440

```

RESULT 7
S53334
streptokinase - Streptococcus sp.
C/Species: Streptococcus sp.
C/Date: 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change 09-Jul-2004
A/Accession: S53334
R/Shi, G.Y.; Chang, B.I.; Chen, S.M.; Wu, D.H.; Wu, H.L.
Biochem. J. 304, 235-241, 1994
A/Title: Function of streptokinase fragments in plasminogen activation.
A/Reference number: S53334; PMID:95091634; PMID:7998939
A/Accession: S53334
A/Status: preliminary
A/Molecule type: protein
A/Residues: 1-197 <SH1>
A/Cross-references: UNIPROT:Q7M115; UNIPARC:UPI000017AC2C

```

Alignment Scores:		Pred. No.:	2,656-46	Length:	197
Score:	788.50	Matches:	193		
Percent Similarity:	47.34%	Conservative:	3		
Best Local Similarity:	46.62%	Mismatches:	1		
Query Match:	25.16%	Indels:	217		
DB:	2	Gaps:	11		

US-09-940-235-11 (1-1782) x S53334 (1-197)

QY	538	ATTGTGAGACCTGAGTGGCTGCTAGAACCGTCCATCTGTCAACCAAGCCAAATGGTGTG	597
DB	1	lIleaIaGIyProGIuItrPleuIeuhAspRyproSerValaSn-----	14
QY	598	AGCGTGTGTAAGTCTGTTGAGGGAGCAATCAAGCATTAAGTCTTAAATTTTGAATTC	657
DB	14	-----	14
QY	658	GATCTAACATCACAGCACTGCTCATATGAGAAAGACAGACAAAGGCTTAAGTCAAAATCA	717
DB	15	-----Ser 15	
QY	718	AAACCAATTTGCTACTGATAGTGGCGCGATGTCACATAAATTGAGAAAAGCTGACTTACTA	777
DB	16	LysProPhealIatInAspSerGIyAlaMetSerHisIySleuGIuIyValaIAspIleu	35
QY	778	AAGGTATTCAGAAACAATTGATCGCTACGTCACAGTAACGACAGACTCTTGAGAGTC	837
DB	36	LysalalIeGIIntrPGIuIleIeu-----	43
QY	838	ATTGATTTTGCAGAGGATGCAACCATTAAGTCGAACGAGCGAAGGTCTACTTGTGCAC	897
DB	44	-----AsnGIyLysValIyPheAlaIAsp 51	
QY	898	AAAGATGGTTCCGTAACTCTTGCCGACCCAACTGTCCAAGATTTTGGTAAAGCGAGAT	957
DB	52	LysAspGIySerValIThr-----	57
QY	958	GTGCGCGCTTAACACCATATTAAGAAAACCAATTAACAAAACGAGAAATCTGTGATGTG	1017
DB	58	-----GluLysProIleGIInsnGIuIAlaIySerValaIAspVal 70	
QY	1018	GAATATACTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTCAGACGAGTCTCAA	1077
DB	71	Glu-----	71
QY	1078	GATATCAAGCTATTGAAAACTAGCTATCGGTGACACCATCACATCTCAAGAAATTACTA	1137
DB	71	-----	71
QY	1138	GCTCAGACCAAAAGCATTTTAAACAAAACGACCGGCTATAGCAATTAAGAACGTAC	1197
DB	72	-----AsnHisProGIyIyThrIleIyGIuIyAsp 82	
QY	1198	TGCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTAAACCAATGGATCAAG	1257
DB	83	SerSerIleVal-----ThrIleIeuProMetAspGIuI 94	
QY	1258	TTTACTTACCGTGTAAATAATCGGAAACAGACTTATAGATCAATTAATAAAATCTGGCTG	1317
DB	95	PheThrIyIy-----SerGIyIeu 100	
QY	1318	AATGAGAAATTAACAACATGACGTGATCTGTGAGAAATTAAGTCTTAAAAAAGG	1377
DB	101	AsnGIuGIuIleAsnValThrAspIleuIleSer-----	111
QY	1378	GAAGAAGCGTATGATCCCTTTGATCGAGTCACTTGAACGTTCACCATCAATTAACGTT	1437
DB	112	-----TyrVal 113	
QY	1438	GATGTGATACCAAGCAATTGCTTAAAAAGTAGACAGCTCTTAACAGTTAGCAAGCTTAAC	1497

Db 114 AspValAspThrAsnGluLeuLeuLysSerGluLeuThrAlaSerGluArgAsn 133
 QY 1498 TTAGACTTCAGAGATTATATAGATCTCGTATAGGCTAAATCTTACAAATCTC 1557
 Db 134 -----AspLeuTyraAspProAspArgAspIleAlaLysLeuLeuTyraAsnAsnLeu 149
 QY 1558 GATGCTTTTGGTATTTATGACTATACCTTAACTGAAAAAGTAGAGATATCAGATGAC 1617
 Db 150 AspAlaPheGlyIleMet----- 155
 QY 1618 ACCAACCGTATCATTAACCGTTTATATGGGACGACCCGAGAGAGAAATGCTAGCTAT 1677
 Db 156 -----IleIleThrValTyraMetGlyLysArgProGluGluAsnAla----- 170
 QY 1678 CATTTAGCCTATGATTAAGATCGTTATATACGAGAAAGAAAGATTACAGTACTG 1737
 Db 171 -----AspArgTyThrGluGluGluGluGluValTyraSerTyLeu 184
 QY 1738 CGTTATACAGGACACCTATATACCTGATTAACCTTACGACAAA 1779
 Db 185 ---TyThrGlyThrProIleProAspAsnProAspAspLys 197

RESULT 8

S77671
 Streptokinase A (EC 3.4.-.-) (alleles 2 and 3) - Streptococcus pyogenes (fragment)
 C:Species: Streptococcus pyogenes
 C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 22-Jun-1999
 C:Accession: S77671; S77672
 R:Kaput, V.; Kanjilal, S.; Hamrick, M.R.; Li, L.L.; Whittam, T.S.; Sawyer, S.A.; Musser, M.L. Microbiol. 16, 509-519, 1995
 A:Title: Molecular population genetic analysis of the streptokinase gene of Streptococcus
 A:Reference number: S77671; MUID:96037795; PMID:7565111
 A:Accession: S77671
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-128 <KAP>
 A:Cross-references: UNIPARC:UPI0000000562; EMBL:U25853; NID:G818908; PIDN:AA85729.1; PI
 A:Experimental source: strain ET1/ML
 A:Note: allele 2
 A:Accession: S77672
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-128 <KAP>
 A:Cross-references: UNIPARC:UPI0000000562; EMBL:U25854; NID:G818910; PIDN:AA85730.1; PI
 A:Experimental source: strain B2/M3
 A:Note: allele 3
 C:Genetics:
 A:Gene: ska
 C:Superfamily: streptokinase
 C:Keywords: hydrolase; plasminogen activator; virulence

Alignment Scores:

Pred. No.: 8.42e-35 Length: 128
 Score: 620.00 Matches: 118
 Percent Similarity: 95.31% Conservative: 4
 Best Local Similarity: 92.19% Mismatches: 6
 Query Match: 19.78% Indels: 0
 Gaps: 0

US-09-940-235-11 (1-1782) x S77671 (1-128)

QY 958 GTGGCGCTTGAACCATTTAAAGAAAAACCAATACAAACGAGGAATCTGTTGATGTG 1017
 Db 1 ValArgValArgProTyrrLysGluLysProValGlnAsnGlnAlaLysSerValAspVal 20
 QY 1018 GAATATACCTGTACAGTTTACCTCTTAAACCTGATGACGATTTCAGACGAGTCTCAA 1077
 Db 21 GluTyrrThrValGlnPheThrProLeuAsnProAspAspPheArgProGluLysLeu 40
 QY 1078 GATACTAAGCTATTGAAAAACATAGCTATCGGTGACACCATCATCTCAAGATTACTA 1137
 Db 41 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 60

QY 1138 GCTCAAGCAACAAGCATTTTAAACAAAAACCAACGAGCTATACGATTATGACGTGAC 1197
 Db 61 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGluTyrrThrIleTyrrGluArgAsp 80
 QY 1198 TCCTCAATCGTCACTGACGACAAATGACATTTTCGTCGATTTTACCAATGATCAAG 1257
 Db 81 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 100
 QY 1258 TTTACTTACCGTGTAAAAATCGGGAACAAGCTTTATAGATCAATAAAAAATCTGTCG 1317
 Db 101 PheThrTyrrHisValLysAsnArgGluGlnAlaTyrrGluIleAsnProLysThrGlyIle 120
 QY 1318 AATGAAGAATATAAACAACACTGAC 1341
 Db 121 LysGluLysThrAsnAsnThrAsp 128

RESULT 9

S77680
 streptokinase A (EC 3.4.-.-) (allele 1) - Streptococcus pyogenes (fragment)
 C:Species: Streptococcus pyogenes
 C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004
 C:Accession: S77680
 R:Kaput, V.; Kanjilal, S.; Hamrick, M.R.; Li, L.L.; Whittam, T.S.; Sawyer, S.A.; Musser, M.L. Microbiol. 16, 509-519, 1995
 A:Title: Molecular population genetic analysis of the streptokinase gene of Streptococcus
 A:Reference number: S77671; MUID:96037795; PMID:7565111
 A:Accession: S77680
 A:Status: translation not shown

A:Molecule type: DNA
 A:Residues: 1-128 <KAP>
 A:Cross-references: UNIPROT:Q54687; UNIPARC:UPI00000BD2D3; EMBL:U25862; NID:G818926; PI
 A:Experimental source: strain ET51/ML7
 C:Genetics:
 A:Gene: ska
 C:Superfamily: streptokinase
 C:Keywords: hydrolase; plasminogen activator; virulence

Alignment Scores:

Pred. No.: 2.59e-31 Length: 128
 Score: 569.00 Matches: 110
 Percent Similarity: 91.41% Conservative: 7
 Best Local Similarity: 85.94% Mismatches: 11
 Query Match: 18.16% Indels: 0
 Gaps: 0

US-09-940-235-11 (1-1782) x S77680 (1-128)

QY 958 GTGGCGCTTGAACCATTTAAAGAAAAACCAATACAAACGAGGAATCTGTTGATGTG 1017
 Db 1 ValArgValArgProTyrrLysGluLysProIleGlnThrProAlaLysSerValAspVal 20
 QY 1018 GAATATACCTGTACAGTTTACCTCTTAAACCTGATGACGATTTCAGACGAGTCTCAA 1077
 Db 21 ArgTyrrThrValGlnPheThrProLeuAsnProAspAspPheArgProValLeuLys 40
 QY 1078 GATACTAAGCTATTGAAAAACATAGCTATCGGTGACACCATCATCTCAAGATTACTA 1137
 Db 41 AsnThrLysLeuLeuLysThrLeuAlaIleGlyLysThrValThrSerGlnGluLeu 60
 QY 1138 GCTCAAGCAACAAGCATTTTAAACAAAAACCAACGAGCTATACGATTATGACGTGAC 1197
 Db 61 AlaGlnAlaGlnSerIleLeuAsnGlnSerHisProAspTyrrThrIleTyrrGluArgAsp 80
 QY 1198 TCCTCAATCGTCACTGACGACAAATGACATTTTCGTCGATTTTACCAATGATCAAG 1257
 Db 81 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 100
 QY 1258 TTTACTTACCGTGTAAAAATCGGGAACAAGCTTTATAGATCAATAAAAAATCTGTCG 1317
 Db 101 PheThrTyrrHisValLysAsnArgGluGlnAlaTyrrGluIleAsnProLysThrGlyIle 120
 QY 1318 AATGAAGAATATAAACAACACTGAC 1341

Db 121 GluGluysThrAsnAsnThrAsp 128

RESULT 10

S77688

streptokinase A (EC 3.4.-.-) (allele 19) - Streptococcus pyogenes (fragment)

C/Species: Streptococcus pyogenes

C/Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004

C/Accession: S77688

R/Kapur, V.; Kanjilal, S.; Hamrick, M.R.; Li, L.L.; Whittam, T.S.; Sawyer, S.A.; Musser, M.O. Microbiol. 16, 509-519, 1995

A/Title: Molecular population genetic analysis of the streptokinase gene of Streptococcus

A/Reference number: S77671; MUID:96037795; PMID:7565111

A/Accession: S77688

A/Status: translation not shown

A/Molecule type: DNA

A/Residues: 1-128 <KAP>

A/Cross-references: UNIPROT:O54695; UNIPARC:UP100000BDBDC; EMBL:U25870; NID:g818942; PIR

A/Experimental source: strain ET56/M72

A/Genetics:

A:Gene: ska

C/Superfamily: streptokinase

C/Keywords: hydrolase; plasminogen activator; virulence

Alignment Scores:

Pred. No.:	1-72e-30	Length:	128
Score:	557.00	Matches:	107
Percent Similarity:	89.84%	Conservative:	8
Best Local Similarity:	83.59%	Mismatches:	13
Query Match:	17.77%	Indels:	0
Gaps:	2	Gaps:	0

US-09-940-235-11 (1-1782) x S77688 (1-128)

Qy 958 GTGGCGGTTAGACCATTAAGAAAAACCAATCAAAACCAAGCAATCTGTGATGTG 1017

Db 1 ValArgValArgProTyrLysGluysProIleGlnThrProAlaLysSerValAspIle 20

Qy 1018 GAATATACCTGTACAGTTTACCTCCCTTAAACCCGATGACGATTTGACAGGTCTCAAA 1077

Db 21 ArgTyrThrValGlnPheThrProLeuAsnProAspAspAspPheLysProValLeuLys 40

Qy 1078 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGATTACTA 1137

Db 41 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrLileThrsSerGlnGluLeuLeu 60

Qy 1138 GCTCAGACCAAGCAATTTAAACAAAAACACCCAGGCTATACGATTTATGACGTCAC 1197

Db 61 AlaGlnAlaGlnSerLileuLileGlnSerLisProAspTyrThrLileTyrGluArgAsp 80

Qy 1198 TCCCTCAATCGTCACATGACATGATTTCCGTCAGATTTTACCAATGATCAAGAG 1257

Db 81 SerSerLileValThrHisAspAsnAspLilePheArgThrLileuProThrAspGlnGlu 100

Qy 1258 TTTACTTACCGTGTAAATAATCGGAAACAAGCTTATAGATCAATAAAAATCTGCTG 1317

Db 101 PheThrTyrHisValLysAsnArgGluGlnAlaTyrLysAlaAsnSerLysThrAspIle 120

Qy 1318 AATGAAGAATAAACAACACTGAC 1341

Db 121 LysGluLysThrAsnAsnThrAsp 128

RESULT 11

S77679

streptokinase A (EC 3.4.-.-) (allele 10) - Streptococcus pyogenes (fragment)

C/Species: Streptococcus pyogenes

C/Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004

C/Accession: S77679

R/Kapur, V.; Kanjilal, S.; Hamrick, M.R.; Li, L.L.; Whittam, T.S.; Sawyer, S.A.; Musser, M.O. Microbiol. 16, 509-519, 1995

A/Title: Molecular population genetic analysis of the streptokinase gene of Streptococcus

A/Reference number: S77671; MUID:96037795; PMID:7565111

A/Accession: S77679

A/Status: translation not shown

A/Molecule type: DNA

A/Residues: 1-128 <KAP>

A/Cross-references: UNIPROT:O54686; UNIPARC:UP100000BDBDF; EMBL:U25861; NID:g818924; PIR

A/Experimental source: strain ET50/M43

A/Genetics:

A:Gene: ska

C/Superfamily: streptokinase

C/Keywords: hydrolase; plasminogen activator; virulence

Alignment Scores:

Pred. No.:	3.77e-30	Length:	128
Score:	552.00	Matches:	106
Percent Similarity:	89.06%	Conservative:	8
Best Local Similarity:	82.81%	Mismatches:	14
Query Match:	17.61%	Indels:	0
Gaps:	2	Gaps:	0

US-09-940-235-11 (1-1782) x S77679 (1-128)

Qy 958 GTGGCGGTTAGACCATTAAGAAAAACCAATCAAAACCAAGCAATCTGTGATGTG 1017

Db 1 ValArgValArgProTyrLysGluysProIleGlnThrProAlaLysSerValAspIle 20

Qy 1018 GAATATACCTGTACAGTTTACCTCCCTTAAACCCGATGACGATTTGACAGGTCTCAAA 1077

Db 21 ArgTyrAlaValGlnPheThrProLeuAsnProAspAspAspPheThrProValLeuLys 40

Qy 1078 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGATTACTA 1137

Db 41 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrLileThrsSerGlnGluLeuLeu 60

Qy 1138 GCTCAGACCAAGCAATTTAAACAAAAACACCCAGGCTATACGATTTATGACGTCAC 1197

Db 61 AlaGlnAlaGlnSerLileuLileGlnSerLisProAspTyrThrLileTyrGluArgAsp 80

Qy 1198 TCCCTCAATCGTCACATGACATGATTTCCGTCAGATTTTACCAATGATCAAGAG 1257

Db 81 SerSerLileValThrHisAspAsnAspLilePheArgThrLileuProThrAspGlnGlu 100

Qy 1258 TTTACTTACCGTGTAAATAATCGGAAACAAGCTTATAGATCAATAAAAATCTGCTG 1317

Db 101 PheThrTyrHisValLysAsnArgGluGlnAlaTyrLysAlaAsnSerLysThrGlyIle 120

Qy 1318 AATGAAGAATAAACAACACTGAC 1341

Db 121 LysGluLysThrAsnAsnThrAsp 128

RESULT 12

FNH1

fibronectin precursor (validated) - human

N/Alternate names: fibronectin splice form ED-A

C/Species: Homo sapiens (hmn)

C/Date: 27-Nov-1985 #sequence_revision 31-Mar-1993 #text_change 31-Dec-2004

C/Accession: A26460; A26284; S03917; A24854; A24476; A91008; A93529; A21011; A90495; A22

R/Dean, D.C.; Bowler, C.L.; Bourgeois, S.

Proc. Natl. Acad. Sci. U.S.A. 84, 1876-1880, 1987

A/Title: Cloning and analysis of the promoter region of the human fibronectin gene.

A/Reference number: A26460; MUID:8715578; PMID:3031656

A/Accession: A26460

A/Molecule type: DNA

A/Residues: 1-49 <DEA>

A/Cross-references: UNIPROT:P02751; UNIPROT:Q14327; UNIPARC:UP1000016A926; GB:M15801; NID

R/Oldberg, A.; Ruoslahti, E.

J. Biol. Chem. 261, 2113-2116, 1986

A/Title: Evolution of the fibronectin gene.

A/Reference number: A26284; MUID:86111901; PMID:3003095

A/Accession: A26284

A/Molecule type: DNA

A/Residues: 1447-1540 <OLD>

A/Cross-references: UNIPARC:UP1000112E37; GB:M12549; NID:g182668

A/Note: the authors translated the codon TTC for residue 1494 as Glu

R/Peollet, G.; Henchcliffe, C.; Sebastio, G.; Baralle, F.E.

Nucleic Acids Res. 16, 3545-3557, 1988

A>Title: Sequence analysis and in vivo expression show that alternative splicing of ED-H
A/Reference number: S00848; PMID:88233940; PMID:3375063
A/Accession: S03917
A/Molecule type: DNA
A/Residues: 1594-1767, 'V', 1769-1783 <PAO>
A/Cross-references: UNIPARC:UPI000017432B; EMBL:X07718; NID:g31402
A/Note: the authors translated the codon AAC for residue 1631 as Asp
R/Vib-Pedersen, K.; Magnusson, S.; Baralle, F.E.
FEBS Lett. 207, 287-291, 1986
A>Title: Donor and acceptor splice signals within an exon of the human fibronectin gene;
A/Reference number: A24854; PMID:87030929; PMID:3770201
A/Accession: A24854
A/Molecule type: DNA
A/Residues: 1992-2147 <VIB>
A/Cross-references: UNIPARC:UPI000017432C; GB:X04530; NID:g31436
R/Gutman, A.; Yamada, K.M.; Kornblith, A.
FEBS Lett. 207, 145-148, 1986
A>Title: Human fibronectin is synthesized as a pre-propolypeptide.
A/Reference number: A24476; PMID:87030890; PMID:3770189
A/Accession: A24476
A/Status: not compared with conceptual translation
A/Molecule type: mRNA
A/Residues: 1-14, 'Q', 16-38 <GUT>
A/Cross-references: UNIPARC:UPI000017432D
R/Kornblith, A.R.; Umezawa, K.; Vib-Pedersen, K.; Baralle, F.E.
EMBO J. 4, 1755-1759, 1985
A>Title: Primary structure of human fibronectin: differential splicing may generate at l
A/Reference number: A91008; PMID:85284965; PMID:2992939
A/Accession: A91008
A/Status: nucleic acid sequence not shown
A/Molecule type: mRNA
A/Residues: 32-1344, 1346-2080; 2112-2386 <KOR>
A/Cross-references: UNIPARC:UPI000017432E; UNIPARC:UPI0000174330; GB:X00739
R/Kornblith, A.R.; Vib-Pedersen, K.; Baralle, F.E.
Nucleic Acids Res. 12, 5853-5868, 1984
A>Title: Human fibronectin: cell specific alternative mRNA splicing generates polypeptid
A/Reference number: A93529; PMID:84272258; PMID:6462919
A/Accession: A93529
A/Molecule type: mRNA
A/Residues: 973-2080; 2112-2386 <KO2>
A/Cross-references: UNIPARC:UPI000017432F; UNIPARC:UPI0000174330; GB:X00739
R/Olberg, A.; Linney, E.; Ruoslahti, E.
J. Biol. Chem. 258, 10193-10196, 1983
A>Title: Molecular cloning and nucleotide sequence of a cDNA clone coding for the cell e
A/Reference number: A21011; PMID:83280929; PMID:6688418
A/Accession: A21011
A/Molecule type: mRNA
A/Residues: 1434-1537 <OL2>
A/Cross-references: UNIPARC:UPI0000174331; GB:K00055; NID:g182680; PIDN:AAAS2459.1; PID
R/Bernard, M.P.; Kolbe, M.; Weil, D.; Chu, M.L.
Biochemistry 24, 2698-2704, 1985
A>Title: Human cellular fibronectin: comparison of the carboxyl-terminal portion with re
A/Reference number: A90495; PMID:85280409; PMID:2992573
A/Accession: A90495
A/Molecule type: mRNA
A/Residues: 1594-2386 <BER>
A/Cross-references: UNIPARC:UPI000016A928; GB:M10905; NID:g182696; PIDN:AAAS2462.1; PID
R/Umezawa, K.; Kornblith, A.R.; Baralle, F.E.
FEBS Lett. 186, 31-34, 1985
A>Title: Isolation and characterization of cDNA clones for human liver fibronectin.
A/Reference number: A22245; PMID:85231203; PMID:2989004
A/Accession: A22245
A/Molecule type: mRNA
A/Residues: 1948-2067 <UME>
A/Cross-references: UNIPARC:UPI000000046A; GB:M27589; NID:g182705; PIDN:AAAS2465.1; PID
A/Accession: B22245
A/Molecule type: mRNA
A/Residues: 1975-1991, 2017-2039 <UM2>
A/Cross-references: UNIPARC:UPI0000174332; UNIPARC:UPI0000174333; GB:M27590
R/Sekiuchi, K.; Kios, A.M.; Kuzachi, K.; Yoshitake, S.; Hakomori, S.
Biochemistry 25, 4936-4941, 1986
A>Title: Human liver fibronectin complementary DNAs: identification of two different mes
A/Reference number: I52394; PMID:87026578; PMID:3021206

A/Accession: I65273
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1978-1990, 2016-2018, 'N', 2020-2081, 2113-2127 <SRK>
A/Cross-references: UNIPARC:UPI00000604C; GB:M14060; NID:g182701; PIDN:AAAS2464.1; PID
R/Kornblith, A.R.; Vib-Pedersen, K.; Baralle, F.E.
Proc. Natl. Acad. Sci. U.S.A. 80, 3218-3222, 1983
A>Title: Isolation and characterization of cDNA clones for human and bovine fibronectin
A/Reference number: A21165; PMID:83221567; PMID:6304699
A/Accession: A21165
A/Molecule type: mRNA
A/Residues: 2291-2386 <KO3>
A/Cross-references: UNIPARC:UPI0000174334; GB:K00799; NID:g182681; PIDN:AAAS2460.1; PID
R/Garcia-Pardo, A.; Pearlstein, E.; Frangione, B.
J. Biol. Chem. 258, 12670-12674, 1983
A>Title: Primary structure of human plasma fibronectin.
A/Reference number: A92398; PMID:84032463; PMID:6630202
A/Accession: A92398
A/Molecule type: protein
A/Residues: 32-47, 'C', 49-51, 'S', 53-72, 'A', 74-290 <GAR1>
A/Cross-references: UNIPARC:UPI0000174335
R/Garcia-Pardo, A.; Gold, L.I.
Arch. Biochem. Biophys. 304, 181-188, 1993
A>Title: Further characterization of the binding of fibronectin to gelatin reveals the
A/Reference number: S34791; PMID:93312001; PMID:8323285
A/Accession: S34791
A/Molecule type: protein
A/Residues: 291-300, 551-560 <GAR2>
A/Cross-references: UNIPARC:UPI0000174336; UNIPARC:UPI0000174337
R/Griffith, C.A.; Calaycay, J.; Shively, J.E.; Smith, R.L.
Thromb. Res. 43, 469-477, 1986
A>Title: Two plasma fibronectin fragments with different gelatin-binding properties.
A/Reference number: A60904; PMID:87019725; PMID:3532418
A/Accession: A60904
A/Molecule type: protein
A/Residues: 293-301 <GRI>
A/Cross-references: UNIPARC:UPI0000174338
R/Calaycay, J.; Pande, H.; Lee, T.; Borsl, L.; Sirl, A.; Shively, J.E.; Zardi, L.
J. Biol. Chem. 260, 12136-12141, 1985
A>Title: Primary structure of a DNA- and heparin-binding domain (domain III) in human p
A/Reference number: A23901; PMID:8608277; PMID:3900070
A/Accession: A23901
A/Molecule type: protein
A/Residues: 616-677, 'Q', 679-703, 'PR', <CAL>
A/Cross-references: UNIPARC:UPI0000174339
R/Pierschbacher, M.D.; Ruoslahti, E.; Sundelin, J.; Lind, P.; Peterson, P.A.
J. Biol. Chem. 257, 9593-9597, 1982
A>Title: The cell attachment domain of fibronectin. Determination of the primary structu
A/Reference number: A92386; PMID:82265604; PMID:7050098
A/Accession: A92386
A/Molecule type: protein
A/Residues: 1441-1548 <PR>
A/Cross-references: UNIPARC:UPI0000141CDS
A/Note: residues 1524-1527 are responsible for the cell-binding activity
R/Garcia-Pardo, A.; Rostagno, A.; Frangione, B.
Biochem. J. 241, 923-928, 1987
A>Title: Primary structure of human plasma fibronectin. Characterization of a 38 kDa do
A/Reference number: A32517; PMID:87241275; PMID:3593230
A/Accession: A32517
A/Molecule type: protein
A/Residues: 1589-1630, 'T', 1722-2058 <GAR3>
A/Cross-references: UNIPARC:UPI000017433A
R/Tresselt, T.; McCarthy, J.B.; Calaycay, J.; Lee, T.D.; Legesse, K.; Shively, J.E.; Pan
Biochem. J. 274, 731-738, 1991
A>Title: Human plasma fibronectin. Demonstration of structural differences between the 2
A/Reference number: S14357; PMID:91190085; PMID:2012601
A/Accession: S14357
A/Molecule type: protein
A/Residues: 1614-1630, 'T', 1722-2081, 2113-2244 <TRE>
A/Cross-references: UNIPARC:UPI000017433B
R/Garcia-Pardo, A.; Pearlstein, E.; Frangione, B.
J. Biol. Chem. 260, 10320-10325, 1985
A>Title: Primary structure of human plasma fibronectin. Characterization of a 31,000-da

A,Reference number: A23891; MUID:85261459; PMID:4019516
A,Accession: A23891
A,Molecule type: protein
A,Residues: 2071-2080/2112-2356 <GAR4>
A,Cross-references: UNIPARC:UPI000017433C; UNIPARC:UPI000017433D
C,Comment: The extra domain and connecting strand 3 are subject to developmental and rib
C,Comment: The cellular and plasma fibronectin are high molecular weight glycoproteins,
action, and transformation.
C,Genetics:
A,Gene: GDB:FNI
A,Cross-references: GDB:119135; OMIM:135600
A,Map position: 2q34-q34
A,Intons: 49/3; 1266/1; 1357/1; 1447/1; 1487/1; 1541/1; 1631/1; 1721/1; 1991/1; 2145/1
C,Superfamily: fibronectin type I repeat homology; fibronectin type II repeat homology;
C,Keywords: acute phase; alternative splicing; cell adhesion; collagen binding; duplicated
F,1-26/Domain: signal sequence #status predicted <SIG>
F,27-31/Domain: propeptide #status predicted <PRO>
F,32-2386/Product: fibronectin #status experimental <MAT>
F,52-87/Domain: fibronectin and heparin binding <FHB>
F,52-272/Domain: fibronectin type I repeat homology <1F1>
F,97-135/Domain: fibronectin type I repeat homology <1F2>
F,141-179/Domain: fibronectin type I repeat homology <1F3>
F,186-225/Domain: fibronectin type I repeat homology <1F4>
F,231-270/Domain: fibronectin type I repeat homology <1F5>
F,308-608/Domain: collagen binding <CBR>
F,308-342/Domain: fibronectin type II repeat homology <1F6>
F,360-401/Domain: fibronectin type II repeat homology <2F1>
F,420-461/Domain: fibronectin type II repeat homology <2F2>
F,470-508/Domain: fibronectin type I repeat homology <1F7>
F,518-555/Domain: fibronectin type I repeat homology <1F8>
F,561-599/Domain: fibronectin type I repeat homology <1F9>

Alignment Scores:
Pred. No.: 1,2e-26 Length: 2386
Score: 503.00 Matches: 101
Percent Similarity: 66.25% Conservative: 5
Best Local Similarity: 63.12% Mismatches: 10
Query Match: 16.05% Indels: 45
DB: 1 Gaps: 3

US-09-940-235-11 (1-1782) x FNHD (1-2386)
QY 254 CCCATAGCTGAGAGAGGTTTGTGATGCTGCGGAGCTTCCATGTCGAGAGAAAG 313
Db 181 ProllealAGLuvCyephephsBAlaAlaGlyThrSeryValValGlyLutrr 200
QY 314 TGGGA-----GAGGCACG 327
Db 201 TrpGluLysProLyrgLnglyTrpMetMetValaAspCysThrCysLeuGlyGluGlySer 220
QY 328 GGAGCATCATTGCACTTCTGAATAGATGCAACGATGAGACACAGACATCTTAT 387
Db 221 GlyArgIleThrCysThrSeraArgAsnAlyGysAsnAspGlnAspThrArgThrSeryr 240
QY 388 AGAATYGGAGACACCTGGAGCAAGAGATATCGAGAAACCTGCTCCAGGTCATCTGC 447
Db 241 ArgIleGlyAspThrTrpSeryrLysAspAsnAlyGlyAsnLeuGlnCysIleCys 260
QY 448 ACAGGCAACGCGCCGAGAGAGTGAAGTGAAGGACACCTCTGTGACACCATG 507
Db 261 ThrGlyAsnGlyArgGlyGlyLutrrLysCysGluAlyGlyHisThrSeraValGlnThrThrSer 280
QY 508 AGCGGATTTGGCCCTTCAACGATGTTGCTATGCT----- 543
Db 281 SerGlySerGlyProPhethrAspValArgAlaAlaValrGlnProGlnProHisPro 300
QY 544 -----GGAAGT 549
Db 301 GlnProProProLyrgLysCysValThrAspSerGlyValValrLyserValGlyMet 320
QY 550 GAGTGGCTGCTAGACCGTTCATCTGTCAACAGCAGCATTTGTTGAGCTTGCTGT 609
Db 321 GlnTrpLeu-----LysThrGlnGlyAsnLysGlnMetLeuGlyThrCysLeuGly 337

RESULT 13
FNBO
fibronectin - bovine
C,Species: Bos primigenius taurus (cattle)
C,Date: 31-Dec-1988 #sequence, revision 31-Dec-1988 #text_change 09-Jul-2004
C,Accession: A26452; B21165; A23292
R,Skortengard, K.; Jensen, M.S.; Sahl, P.; Petersen, T.B.; Magnusson, S.
Eur. J. Biochem. 161, 441-453, 1986
A,Title: Complete primary structure of bovine plasma fibronectin.
A,Reference number: A26452; MUID:87054047; PMID:3780752
A,Accession: A26452
A,Molecule type: protein
A,Residues: 1-2265 <SKO>
A,Cross-references: UNIPROT:P07589; UNIPARC:UPI000012A7BE
R,Kornilic, A.R.; Vilde-Pedersen, K.; Baralle, F.E.
Proc. Natl. Acad. Sci. U.S.A. 80, 3218-3222, 1983
A,Title: Isolation and characterization of cDNA clones for human and bovine fibronectin
A,Reference number: A21165; MUID:83221567; PMID:6304699
A,Accession: B21165
A,Molecule type: mRNA
A,Residues: 2170-2265 <KOR>
A,Cross-references: UNIPARC:UPI000016C30C; GB:X00800; NID:G163055; PIDN:AAA30521.2; PID:
R,Petersen, T.B.; Thogersen, H.C.; Skortengard, K.; Vilde-Pedersen, K.; Sahl, P.; Sottr
Proc. Natl. Acad. Sci. U.S.A. 80, 137-141, 1983
A,Title: Partial primary structure of bovine plasma fibronectin: three types of internal
A,Reference number: A23292; MUID:83117805; PMID:6218503
A,Accession: A23292
A,Molecule type: protein
A,Residues: 1-16, 'C', 18-20, 'S', 22-43/447-463; 1367-1517; 1567-1673; 2062-2176, 'N', 2178-2265
A,Cross-references: UNIPARC:UPI000017433E; UNIPARC:UPI000017433F; UNIPARC:UPI0000174340;
C,Comment: The plasma fibronectin molecule consists of two chains, which are connected b
C,Comment: Fibronectins bind cell surfaces and various compounds including collagen, fib
aling, and maintenance of cell shape.
C,Comment: Plasma fibronectin is synthesized by hepatocytes.
C,Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II rep
C,Keywords: acute phase; alternative splicing; collagen binding; duplication; extracellular
F,21-241/Domain: fibrin and heparin binding <FBR>
F,21-56/Domain: fibronectin type I repeat homology <1F1>
F,66-104/Domain: fibronectin type I repeat homology <1F2>
F,110-148/Domain: fibronectin type I repeat homology <1F3>
F,155-194/Domain: fibronectin type I repeat homology <1F4>
F,200-239/Domain: fibronectin type I repeat homology <1F5>
F,277-577/Domain: collagen binding <CBR>
F,389-430/Domain: fibronectin type I repeat homology <1F6>
F,329-370/Domain: fibronectin type II repeat homology <2F2>
F,439-477/Domain: fibronectin type II repeat homology <2F3>
F,487-524/Domain: fibronectin type I repeat homology <1F8>
F,530-568/Domain: fibronectin type I repeat homology <1F9>
F,578-661/Domain: fibronectin type II repeat homology <FN3A>
F,688-770/Domain: fibronectin type III repeat homology <FN3B>
F,779-860/Domain: fibronectin type III repeat homology <FN3C>
F,875-957/Domain: fibronectin type III repeat homology <FN3D>
F,965-1046/Domain: fibronectin type III repeat homology <FN3E>
F,1055-1134/Domain: fibronectin type III repeat homology <FN3F>
F,1142-1227/Domain: fibronectin type III repeat homology <FN3G>
F,1235-1318/Domain: fibronectin type III repeat homology <FN3H>
F,1326-1404/Domain: fibronectin type III repeat homology <GN3I>
F,1410-1517/Domain: cell attachment <CAD>
F,1416-1502/Domain: fibronectin type III repeat homology <FN3J>
F,1493-1495/Region: cell attachment (R-G-D) motif
F,1510-1592/Domain: fibronectin type III repeat homology <FN3K>
F,1600-1870/Domain: heparin binding <HB2>
F,1600-1682/Domain: fibronectin type III repeat homology <FN3L>
F,1692-1773/Domain: fibronectin type III repeat homology <FN3M>
F,1781-1863/Domain: fibronectin type III repeat homology <FN3N>
F,1970-1972/Region: cell attachment (R-G-D) motif
F,1982-2062/Domain: fibronectin type III repeat homology <FN3O>
F,1985-2216/Domain: fibronectin type III repeat homology <FN3P>
F,2085-2124/Domain: fibronectin type I repeat homology <1F10>
F,2130-2167/Domain: fibronectin type I repeat homology <1F11>

F/2174-2209/Domain: fibronectin type I repeat homology <1F12>
F/1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F/3/Cross-link: isopeptide (Gln) (interchain to Lys N6-amino of fibrin) #status experime
F/21-47/45-56/66-94/92-104/110-138/136-148/155-184/182-199/200-229/227-239/277-304/302-3
7/2155-2167/2174-2200/2198-2209/disulfide bonds: #status predicted
F/1399/497/511/846/976/1213/1987/Binding site: carbohydrate (Asn) (covalent) #status expe
F/1205/1692/Binding site: carbohydrate (Asn) (covalent) #status absent
F/1943/1944/Binding site: carbohydrate (Thr) (covalent) #status experimental
F/2246/Disulfide bonds: interchain (to 2250) #status predicted
F/2250/Disulfide bonds: interchain (to 2246) #status predicted
F/2263/Binding site: phosphate (Ser) (covalent) #status experimental

Alignment Scores:
Pred. No.: 3,8e-25 Length: 2265
Score: 481.00 Matches: 97
Percent Similarity: 65.00% Conservative: 7
Best Local Similarity: 60.62% Mismatches: 12
Query Match: 15.35% Indels: 45
DB: 1 Gaps: 3

US-09-940-235-11 (1-1782) x FNBO (1-2265)

QY	254	CCCATAGCTGAGAGAGCTTTTATCATGCTGCTGGAGACTTCTTATGTTGTCGAGAAACG	313
DB	150	ProilaalaglulysCysPheaspGlnAlaGlyThrSerYrValIglYglutThr	169
QY	314	TGGGA-----	GAGGAGC 327
DB	170	TTPglulysProTyrGlnGlyTTPMetMetValaspCysThrCysleuGlyGlnGlySer	189
QY	328	GGAGCATCACTTGCACTTCTAGAAATAGATGACAGCATGACAGACATCTTAT	387
DB	190	GlyAyrGIllethrcysThrSerAraGlnAraGlyAspAsnaspGlnAraGlnThrSerTyr	209
QY	388	AGAATTGAGACACCTTGAGCAAGAAAGATATGAGAAACCTGCTCCAGTGCATTC	447
DB	210	ArgGlllGlyAspThrTTPSerTylsAspAsnAraGlyAsnleuLeuGlnCysIleCys	229
QY	448	ACAGGAGAGCGCCGAGAGAGTGGAGTGGAGAGCACCCTGCTGGAGACCATTCG	507
DB	230	ThnglyAsnGlyAraGlyGlyutTTPYsCysGlnAraGlnThrSerleuGlnThrThrSer	249
QY	508	AGCGGATTCGCGCCCTTCACCGATGCTTCGATTCCT-----	543
DB	250	AlaGlySerGlySerTherThraPvalAraGlnThrAlaIleTyrGlnProGlnProHisPro	269
QY	544	-----	GAGCT 549
DB	270	GlnProProTyrGlyGlnHisCysValThraPserGlyValValTyrSerValGlyMet	289
QY	550	GAGTGGCTGCTAGACCGCTCATCTGTCACACAGCAGCATTTGGTTTACGCTTGGT	609
DB	290	GlnTTPleu-----	lysthrGlnGlyAsnlyslmetleuCysThrCysleuGly 306

RESULT 14
S14428
fibronectin precursor - rat
C/Species: Rattus norvegicus (Norway rat)
C/Date: 28-Oct-1994 #sequence revision 28-Oct-1994 #text change 09-Jul-2004
C/Accession: S14428; S12455; A22319; S46203; S00459; A27252; I59049
R/Hynes, R.O.
submitted to the EMBL Data Library, July 1989
A/Reference number: S14428
A/Accession: S14428
A/Molecule type: mRNA
A/Residues: 1-2477 <HYN>
A/Cross-references: UNIPROT:P04937; UNIPARC:UPI000012A7C6; EMBL:X15906; NID:G56163; PID:
R.Schwarzbauer, U.E.; Patel, R.S.; Fonda, D.; Hynes, R.O.
EMBO J. 6, 2573-2580, 1987
A/Title: Multiple sites of alternative splicing of the rat fibronectin gene transcript.
A/Reference number: S12455; MUID:88054951; PMID:2445560
A/Accession: S12455
A/Status: nucleic acid sequence not shown

A/Molecule type: mRNA
A/Residues: 609-1810, 'T', 1812-2283 <SCH>
A/Cross-references: UNIPARC:UPI0000177AE; EMBL:X15906
R/Tamkun, J.W.; Schwarzbauer, J.E.; Hynes, R.O.
Proc. Natl. Acad. Sci. U.S.A. 81, 5140-5144, 1984
A/Title: A single rat fibronectin gene generates three different mRNAs by alternative s
A/Reference number: A22319; MUID:84298097; PMID:6089177
A/Accession: A22319
A/Molecule type: DNA
A/Residues: 2052-2237 <TAM>
A/Cross-references: UNIPARC:UPI0000177ABE
R/Falkenberg, C.; Enghild, J.J.; Thøgersen, I.B.; Salvesen, G.; Akerstrøm, B.
Biochem. J. 301, 745-751, 1994
A/Title: Isolation and characterization of fibronectin-alpha(1)-microglobulin complex
A/Reference number: S46203; MUID:94330948; PMID:7519849
A/Accession: S46203
A/Status: Preliminary
A/Molecule type: protein
A/Residues: 1183-1192, 'Gln', 1268, 'P', 1270-1271, 'D', 1273, 'CF', 1276, 'PY', 1385-1399 <FAL>
A/Cross-references: UNIPARC:UPI0000177AEC; UNIPARC:UPI0000177AED; UNIPARC:UPI0000177AEE
R/Patel, R.S.; Odermatt, E.; Schwarzbauer, J.E.; Hynes, R.O.
EMBO J. 6, 2565-2572, 1987
A/Title: Organization of the fibronectin gene provides evidence for exon shuffling duri
A/Reference number: S00459; MUID:88054950; PMID:3119323
A/Accession: S00459
A/Molecule type: DNA
A/Residues: 1-139/2382-2477 <PAT>
A/Cross-references: UNIPARC:UPI000017095C; UNIPARC:UPI0000177AEF; EMBL:X05831
A/Note: the authors translated the codon CCT for residues 51 and 94 as Ala
R/Schwarzbauer, J.E.; Tamkun, J.W.; Lemieche, I.R.; Hynes, R.O.
Cell 35, 421-431, 1983
A/Title: Three different fibronectin mRNAs arise by alternative splicing within the cod
A/Reference number: A27252; MUID:84082067; PMID:6317187
A/Accession: A27252
A/Molecule type: mRNA
A/Residues: 1586-1720, 'T', 1722, 1813-2477 <SC2>
A/Cross-references: UNIPARC:UPI0000177AF0
R/Odermatt, E.; Tamkun, J.W.; Hynes, R.O.
Proc. Natl. Acad. Sci. U.S.A. 82, 6571-6575, 1985
A/Title: Repeating modular structure of the fibronectin gene: Relationship to protein s
A/Reference number: I59049; MUID:86016741; PMID:3863113
A/Accession: I59049
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1722-1810 <RBS>
A/Cross-references: UNIPARC:UPI00000040C; GB:M11750; NID:G204164; PID:AAA41170.1; PID
C/Genetics:
A/Insertions: 51/1; 94/1; 2416/3; 2454/3
C/Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II rep
C/Keywords: alternative splicing; cell adhesion; collagen binding; disulfide bond; dupl
F/1-32/Domain: signal sequence #status predicted <SIG>
F/33-2477/Product: fibronectin #status predicted <MAT>
F/53-88/Domain: fibronectin type I repeat homology <1F1>
F/98-136/Domain: fibronectin type I repeat homology <1F2>
F/142-180/Domain: fibronectin type I repeat homology <1F3>
F/187-226/Domain: fibronectin type I repeat homology <1F4>
F/232-271/Domain: fibronectin type I repeat homology <1F5>
F/308-342/Domain: fibronectin type I repeat homology <1F6>
F/360-401/Domain: fibronectin type II repeat homology <2F1>
F/420-461/Domain: fibronectin type II repeat homology <2F2>
F/470-508/Domain: fibronectin type I repeat homology <1F7>
F/518-555/Domain: fibronectin type I repeat homology <1F8>
F/561-599/Domain: fibronectin type I repeat homology <1F9>
F/609-692/Domain: fibronectin type III repeat homology <FN3A>
F/718-800/Domain: fibronectin type III repeat homology <FN3B>
F/809-890/Domain: fibronectin type III repeat homology <FN3C>
F/905-967/Domain: fibronectin type III repeat homology <FN3D>
F/995-1076/Domain: fibronectin type III repeat homology <FN3E>
F/1085-1164/Domain: fibronectin type III repeat homology <FN3F>
F/1172-1257/Domain: fibronectin type III repeat homology <FN3G>
F/1265-1348/Domain: fibronectin type III repeat homology <FN3H>
F/1356-1439/Domain: fibronectin type III repeat homology <FN3I>
F/1447-1529/Domain: fibronectin type III repeat homology <FN3J>

F1537-1619/Domain: fibronectin type III repeat homology <FN3X>
 F1614-1616/Region: cell attachment (R-G-D) motif
 F1631-1713/Domain: fibronectin type III repeat homology <FN3L>
 F1721-1803/Domain: fibronectin type III repeat homology <FN3M>
 F1811-1893/Domain: fibronectin type III repeat homology <FN3O>
 F1903-1984/Domain: fibronectin type III repeat homology <FN3O>
 F1992-2074/Domain: fibronectin type III repeat homology <FN3O>
 F2181-2183/Region: cell attachment (R-G-D) motif
 F2193-2273/Domain: fibronectin type III repeat homology <FN3Q>
 F2226-2235/Domain: fibronectin type I repeat homology <FI10>
 F2341-2378/Domain: fibronectin type I repeat homology <FI11>
 F2385-2420/Domain: fibronectin type I repeat homology <FI12>
 F253-79, 77-88, 98-124, 124-136, 142-170, 168-180, 187-216, 214-226, 232-261, 259-271, 308-335, 333
 F2458/Disulfide bonds: interchain (to 2462) #status predicted
 F2462/Disulfide bonds: interchain (to 2458) #status predicted

Alignment Scores:

Pred. No.:	2,5e-23	Length:	2477
Score:	454.50	Matches:	92
Percent Similarity:	63.75%	Conservative:	10
Best Local Similarity:	57.50%	Mismatches:	13
Query Match:	14.50%	Indels:	46
DB:	2	Gaps:	4

US-09-940-235-11 (1-1782) x S14428 (1-2477)

```

OY 254 CCCATAGCTGAGAGAGTGTGATCATGCTGCGACTTCTTANGTGTCGGAGAAACG 313
Db 182 ProtleaIaGluIyCySpheAspHisAlaIaGlyThrSerTyValIaGlyIuThr 201
OY 314 TGGGA-----GAGGCAGC 327
Db 202 TrpGluIySPProTyGlnGlyTrpMetMetValaSPCySerThrCysLeuGlyGluIyAaH 221
OY 328 GAGAGCATCACTTGCACTTCTAGAAATAGATGCAACATGACAGACACAAAGACATCTAT 387
Db 222 GlyAIGIleThrCySerTherArgaenArgCySaenAspGlnSpThrArgThrSerTyR 241
OY 388 AGAATTGAGACACCTGAGAGCAAGAGATTAATCGAGAAACCTGCTCCAGTGCATCTGC 447
Db 242 ArgIleGlyAspThrTrpSerIySlySaAspAsnArgGlyAaenLeuGlnCysValCys 261
OY 448 ACAAGCAACGCCCGAGAGAGAGTGAAGTGTGAGAGGACACCTCTGTGACACACATCG 507
Db 262 ThrGlyAsnGlyArgGlyGluTrpIyS CysGluArgHisVal---LeuGlnSerAlaSer 280
OY 508 AGCGGATCGGCCCTTCACCGATGTCGATTGCT----- 543
Db 281 AlaGlySerGlySerPheThrAspValaArgThrAlaIleTyGlnProGlnThrHisPro 300
OY 544 -----GGACCT 549
Db 301 GlnProAlaProTyGlyHisCysValThrAspSerGlyValValTySerValGlyMet 320
OY 550 GAGTGGCTGCTAGACCGTTCATCTGTCAACAACGCCAATTGGTGTAGCGTGTCTGCT 609
Db 321 GlnTrpLeu-----LysSerGlnGlyAspIySglnMetLeuCySerThrCysLeuGly 337

```

RESULT 15

S77676

streptokinase A (EC 3.4.-.-) (allele 7) - Streptococcus pyogenes (fragment)

C1Species: Streptococcus pyogenes

C1Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004

C1Accession: S77676

R1Kaput, V.; Kanjilal, S.; Hamrick, M.R.; Li, L.L.; Whittam, T.S.; Sawyer, S.A.; Musser, Mol. Microbiol. 16, 509-519, 1995

A1Title: Molecular population genetic analysis of the streptokinase gene of Streptococcus

A1Reference number: S77671; MUID:96037795; PMID:7565111

A1Accession: S77676

A1Status: translation not shown

A1Molecule type: DNA

A1Residues: 1-128 <KAP>

A1Cross-references: UNIPROT:O54683; UNIPARC:UPI0000087143; EMBL:U25858; NID:G818918; PDB:
 A1Experimental source: ETE/M66
 A1Genetics:
 A1Gene: SKA
 C1Superfamily: streptokinase
 C1Keywords: hydrolase; plasminogen activator; virulence

Alignment Scores:

Pred. No.:	2,51e-21	Length:	128
Score:	423.00	Matches:	85
Percent Similarity:	75.78%	Conservative:	12
Best Local Similarity:	66.41%	Mismatches:	31
Query Match:	13.50%	Indels:	0
DB:	2	Gaps:	0

US-09-940-235-11 (1-1782) x S77676 (1-128)

```

OY 958 GTGCGGCTTACAGACATTAAGAAACCAATACAAACCAAGCAAAATCTGTGATGTG 1017
Db 1 ValArgValArgProTyGlnProIySAlaValHisAsnSerAlaGluArgValaSerVal 20
OY 1018 GAATPACTGTACAGTTTACTCCCTTAACCTGTAGACGATTTGAGACAGGCTTCAAA 1077
Db 21 AsnTyGluValSerPheValSerGluThrGlyAspLeuAspPheThrProSerLeuArg 40
OY 1078 GATACCTAAGCTATTGAAAACACTAGCTATCGTGACACCATCAATCTCAAGATTACTA 1137
Db 41 AspArgTyHisIleuThrThrIleuAlaValaGlyAspSerIleuSerGlnIleuAla 60
OY 1138 GCTCAAGCAACAAAGCATTTTAAACAAACCAACCGAGGCTATGATTTATGAAACGTGAC 1197
Db 61 AlaIleAlaGlnPheIleuSerIySglnHisProAspTyRlleIleThrIySArgAsp 80
OY 1198 TCCTCAATGCTCATCATGACAAATGACATTTTCGTACGATTTTACCAATGATCAAG 1257
Db 81 SerSerIleValThrHisAspAsnAspIlePheArgThrIleuProMetAspGlnGlu 100
OY 1258 TTTACTTACCGGTTTAAATCGGAAACAAAGCTTATAGATCAATATAAATCTGCTGCTG 1317
Db 101 PheThrTyHisValIySAsnArgGlnIleAlaTyGlyIleAsnIySlySerGlyGln 120
OY 1318 AATGAAGAAATTAACAACCTGAC 1341
Db 121 LysGluIySleAsnAsnThrAsp 128

```

Search completed: January 28, 2006, 02:37:54
 Job time : 44.3089 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Comugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 28, 2006, 01:48:09 ; Search time 192.043 Seconds
(without alignments)
13093.422 Million cell updates/sec

Title: US-09-940-235-11

Perfect score: 3134
Sequence: 1 tcgcctcacgctcgtcgcg.....ataacctaacgacaataa 1782

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2166443 segs, 705528306 residues

Total number of hits satisfying chosen parameters: 4332886

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

MODEL=frame+ n2p model -DEV=rlp
Q=/sgn2.1/USPTO_epool_p/US09940235/runat_27012006_144218_27563/app.query.fasta_1.7708
-DB=UniProt -QFMT=fastan -SUFFIX=rup -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=DCT -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09940235 @CGN_1.1.1355 @runat_27012006_144218_27563 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEUSER -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : UniProt 05.80.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2158	68.9	440	1 STRP_STREQ	P00779 streptococc
2	2129	67.9	414	2 Q53284_STREQ	Q53284 streptococc
3	2117	67.5	436	3 Q6UK57_STREQ	Q6UK57 streptococc
4	2116	67.5	440	1 STRP_STRS1	P10519 streptococc
5	1949	62.2	440	2 Q7X0V1_STRPY	Q7X0V1 streptococc
6	1937	61.8	440	1 STRP_STRPY	P10520 streptococc
7	1937	61.8	440	2 Q53XK6_STRPY	Q53XK6 streptococc
8	1937	61.8	440	2 Q8K5R8_STRP3	Q8K5R8 streptococc
9	1930	61.6	440	2 Q7X0Y2_STRPY	Q7X0Y2 streptococc
10	1929	61.6	440	2 Q7X0Y8_STRPY	Q7X0Y8 streptococc
11	1926	61.5	440	2 Q7X0Y7_STRPY	Q7X0Y7 streptococc
12	1926	61.5	440	2 Q5X9T6_STRP6	Q5X9T6 streptococc
13	1923	61.4	440	2 Q8NZA6_STRP8	Q8NZA6 streptococc
14	1880	60.0	440	2 Q7X0Y3_STRPY	Q7X0Y3 streptococc
15	1872	59.7	440	2 Q7X0Y0_STRPY	Q7X0Y0 streptococc
16	1866	59.5	440	2 Q7X0Y5_STRPY	Q7X0Y5 streptococc

17	1865	59.5	440	2 Q7X0X8_STRPY	Q7X0X8 streptococc
18	1847	58.9	440	2 Q7X0X9_STRPY	Q7X0X9 streptococc
19	1845	58.9	440	2 Q7X0X7_STRPY	Q7X0X7 streptococc
20	1838	58.6	440	2 Q57391_STRPY	Q57391 streptococc
21	1825	58.2	440	2 Q7X0Y6_STRPY	Q7X0Y6 streptococc
22	1810	57.8	440	1 STRQ_STRPY	P96471 streptococc
23	1792	57.2	432	2 Q7X0Y4_STRPY	Q7X0Y4 streptococc
24	788.5	25.2	197	2 Q7M115_STRP	Q7M115 streptococc
25	728	23.2	141	2 Q7X0X2_STREQ	Q7X0X2 streptococc
26	727	23.2	141	2 Q7X0X3_STREQ	Q7X0X3 streptococc
27	724	23.1	141	2 Q7X0W1_STREQ	Q7X0W1 streptococc
28	723	23.1	141	2 Q7WS87_STREQ	Q7WS87 streptococc
29	717	22.9	141	2 Q7X0W3_STREQ	Q7X0W3 streptococc
30	717	22.9	141	2 Q7X0X4_STREQ	Q7X0X4 streptococc
31	715	22.8	141	2 Q7X0W5_STREQ	Q7X0W5 streptococc
32	714	22.8	141	2 Q7X0X5_STREQ	Q7X0X5 streptococc
33	712	22.7	141	2 Q7X0X6_STREQ	Q7X0X6 streptococc
34	711	22.7	141	2 Q7X0W2_STREQ	Q7X0W2 streptococc
35	711	22.7	141	2 Q7X0W9_STREQ	Q7X0W9 streptococc
36	710	22.7	141	2 Q7X0M6_STREQ	Q7X0M6 streptococc
37	707	22.6	141	2 Q7X0W0_STREQ	Q7X0W0 streptococc
38	705	22.5	141	2 Q7X0V5_STRPY	Q7X0V5 streptococc
39	704	22.5	141	2 Q7X0X0_STREQ	Q7X0X0 streptococc
40	701	22.4	141	2 Q7X0W7_STREQ	Q7X0W7 streptococc
41	697	22.2	141	2 Q7X0W4_STREQ	Q7X0W4 streptococc
42	696	22.2	141	2 Q7X0M8_STREQ	Q7X0M8 streptococc
43	693	22.1	141	2 Q7X0X1_STREQ	Q7X0X1 streptococc
44	690	22.0	141	2 Q7X0R2_STRPY	Q7X0R2 streptococc
45	686	21.9	141	2 Q7X0S4_STRPY	Q7X0S4 streptococc

ALIGNMENTS

RESULT 1
STRP_STREQ STRP_STREQ STANDARD; PRT; 440 AA.
AC P00779;
DT 21-Jul-1986 (Rel. 01, Created)
DT 01-Apr-1988 (Rel. 07, Last sequence update)
DT 10-May-2005 (Rel. 47, Last annotation update)
DB Streptokinase C precursor.
GN Name=skc;
OS Streptococcus equisimilis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=119602;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=H46A;
RX MEDLINE=85232082; PubMed=2989113; DOI=10.1016/0378-1119(85)90145-3;
RA Walke H., Roe B.A., Ferretti J.U.;
RT "Nucleotide sequence of the streptokinase gene from Streptococcus
equisimilis H46A.";
RL Gene 34:357-362(1985).
RN [2]
RP PROTEIN SEQUENCE OF 27-440.
RX MEDLINE=83127125; PubMed=6760891;
RA Jackson K.W., Tang J.;
RT "Complete amino acid sequence of streptokinase and its homology with
serine proteases.";
RL Biochemistry 21:6620-6625(1982).
-!- FUNCTION: This protein is not a protease, but it activates
plasminogen by complexing with it. As a potential virulence
factor, it is thought to prevent the formation of effective fibrin
barriers around the site of infection, thereby contributing to the
invasiveness of the cells.

This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.

```
DR EMBL; K02986; AAA26974.1; -; Genomic DNA.
DR EMBL; X72832; CAA51351.1; -; Genomic DNA.
DR PIR; A00967; BZSO.
DR PIR; A22801; A22801.
DR PDB; 1BML; X-ray; C/D=38-399.
DR PDB; 1L4D; X-ray; B=40-173.
DR PDB; 1L4Z; X-ray; B=27-173.
DR PDB; 1OOR; X-ray; A/B/C/D=177-314.
DR SMR; P00779; 38-398.
DR InterPro; IPR004093; Staphylokinase.
DR InterPro; IPR008124; Streptokinase.
DR Pfam; PF02821; Staphylokinase; 3.
DR PRINTS; PR01753; STREPKINASE.
DR 3D-structure; Direct protein sequencing; Plasminogen activation;
KW Signal; Virulence.
FT SIGNAL 1 26
FT CHAIN 27 440 Streptokinase C.
FT VARIANT 195 195 L -> D.
FT VARIANT 207 207 D -> L.
FT CONFLICT 298 300 EKY -> LEYK (in Ref. 2).
FT CONFLICT 438 438 N -> D (in Ref. 2).
FT STRAND 180 180
FT STRAND 184 194
FT TURN 199 200
FT TURN 203 204
FT STRAND 205 205
FT STRAND 209 214
FT TURN 216 217
FT STRAND 219 221
FT HELIX 222 236
FT TURN 238 239
FT STRAND 240 252
FT TURN 253 254
FT STRAND 259 260
FT STRAND 268 270
FT STRAND 275 275
FT STRAND 278 280
FT TURN 282 284
FT STRAND 287 289
FT STRAND 292 304
FT TURN 305 306
SQ SEQUENCE 440 AA; 50140 MW; 8FC1F22648AC77A CRC64;

Alignment Scores:
Pred. No.: 7.63e-142 Length: 440
Score: 2158.00 Matches: 414
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 68.86% Indels: 0
DB: 1 Gaps: 0

US-09-940-235-11 (1-1782) x STRP_STRBQ (1-440)
QY 538 ATTGCTGACCTGATGCTGCTAGACGCTCATCTGTCAACAACGCAATTTGTT 597
DB 27 IleaaglyPProgluITripLeuunspatgProseValasmbnsersglneuvai 46
QY 598 AGCGTGTGCTACTGTGAGGGAAGATCAAGATTAGTCTTAATTTTGAATC 657
DB 47 SerValaIaaglyThrValgluglyThraanglInaspIleSerleuysrhepgeIuile 66
QY 658 GATTTAACATCCAGACCTGCTCATGAGGAAAGACAGACAGAGCTTAAGTCCAAATCA 717
DB 67 AspleuITrSerArgrProIahIsaglyLystrGlInglInlyeuserProlysser 86
QY 718 AAACATTGTGCTAGTATGAGCGGATGTCACATTAACCTTGAAGAGCTGACTACTA 777
DB 87 LysPProhealathrAspsersglYalawetSerHslyLeucluuYsaIaaspleu 106
QY 778 AAGGCTATTCAAGAACATTTGATCGCTAACGTCACAGTAAACGACGACTTATTTGAGTC 837
DB 107 LysaIaIeIeIngInglInleuIeIaIaenValHIsSerAspAspAspTyrrhegluVal 126
```

```
QY 838 ATTGATTTTGCAGCGATGCAACCATTAATGATGCAAAACGCAAGTCTACTTGTCTGAC 897
DB 127 IleaSPhealAserAspIathrIleThraSPargnsnIyusValTyrrhealAsp 146
QY 898 AAAAGATGTTCCGTTAACTTTGCCGACCCAACTGCTCCAGAAATTTTGTAAACGCAAT 957
DB 147 LysAspOlYserValThrlleuProthInglInProValglInglInPheleuSerGlYHIs 166
QY 958 GTCGCGCTTACCATTAATTAAGAAAACCAATCAAAACCAAGCAAAATCTTTGAAGTG 1017
DB 167 ValArgValArgProTyrrusgluYsProIleGlInasnlalYsSerValAspVal 186
QY 1018 GAATATPACTGATGATTTATCCCTTAAACCTGATGACGATTTCAACAGGCTCAAA 1077
DB 187 GluTyrrThrValglInPheThrrProleuAsnProAspAspAspPhehArgProglYleuYls 206
QY 1078 GATPACTAAGCTATTGAAAAACATAGCTATGCGTGAACCATCATCTCAAGAAATTACTA 1137
DB 207 AspThrlyleuYsleuYsThrlleuAlaIleglYAspThrlleThrsersglInglInleu 226
QY 1138 GCTCAAGCACAAGACATTTTAAACAAAACCCAGGCTATACGATTTATGAACGTGAC 1197
DB 227 AlaGlInaIaIeInserIleuAsnlySAsnHIsProglYTyrrThrlleTyrgluArgAsp 246
QY 1198 TCCCTCAATGCTGACTATGACATATGACATTTCCGTCGATTTTACCAATGATCAAGAG 1257
DB 247 SerSerIleValThrHIsAspAsnAspIlePheArgThrlleuProMetAspGInglIn 266
QY 1258 TTTACTTACCGTGTAAAAATCCGGAACAAAGCTTATAGATCAATAAAAATCGTCTG 1317
DB 267 PheThrrTyrrglYalYlsAsnArgglInglInalatyrrglYleasnlySAspSerGlYleu 286
QY 1318 AATGAAGAATTAACAACACTGACCTGATCTCGAAGAAATTTACGCTTTAAAAAGG 1377
DB 287 AsnGlugluInleAsnAsnThraSPleuIleSerGlYusTyrrTyrrValleuYslyeGlY 306
QY 1378 GAAAGCGGTATGATCCCTTGTATGCGACACTGTGAACCTGTACACATCAATATAGCTT 1437
DB 307 GluYsProTyrrAspProPheAspArgSerHIsleuYsleuPheThrlleYsTyrrVal 326
QY 1438 GATGTCATCAACCAAGATTTGCTAAAAAGTGAACAGCTCTTAAACGCTACGACGATTAAC 1497
DB 327 AspValAspThrsnIeInglInleuYsSerSerGlugluInleuThraIAsersgluArgAsn 346
QY 1498 TTAGACTTCAGAGATTTATTCGATCTCTGATTAAGGCTTAACTCTTACACATCTC 1557
DB 347 LeuAspPheArgAspLeuTyrrAspProArgAspLyseIalYsleuYsleuTyrrAsnAsnleu 366
QY 1558 GATGCTTTGGTATTTATGACATTAACCTTAACTGGAAGAAAGTAAAGATTAATCCAGTAC 1617
DB 367 AspAlaPheglYIleMetAspTyrrThrlleuThrglyValgluAspAsnHIsAspAsp 386
QY 1618 ACCAACGCTATCAATACCGTTTATATGGGCAACGACCCGAAAGAGAGAAATGCTAGCTAT 1677
DB 387 ThrAsnArgIleIethrValTyrrMetglYlyAspArgProIuInglInuAsnIaIAserTyrr 406
QY 1678 GATTTAAGCTTATTAAGATGCTTATACCGAAGAAAGACGAGAAAGTTTACGCTACTG 1737
DB 407 HIsleuAlatyrrAspysAspArgTyrrThrgluInglIngluArgluValTyrrSerTyrrleu 426
QY 1738 CGTTTATACAGGACACCTTATACCTGATTAACCTTAAGACAAA 1779
DB 427 ArgTyrrThrglyThrrProIleProAspAsnProAsnAspLyS 440

RESULT 2
ID Q53284_STRBQ PRELIMINARY; PRT; 414 AA.
AC Q53284;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE SKC-2.
OS Streptococcus equisimilis.
```

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_Taxid=119602;
 RN (1)
 RP NUCLEOTIDE SEQUENCE.
 RA Estrada M.P., Hernandez L., Perez A., Rodriguez P., Serrano R.,
 RA Rubiera R., Pedraza A., Padron G., Antuch W., la Fuente J., et al.;
 RT "High level expression of streptokinase in *Escherichia coli*.";
 RL Biotechnology 0:1138-1142(1992).
 DR EMBL; S46536; AAC60418.1; -; Genomic_DNA.
 DR PDB; 1C4P; X-ray; A/B/C/D=149-285.
 DR SMR; 053284; 12-372.
 DR GO; GO:0008243; F:plasmaeogen activator activity; IEA.
 DR GO; GO:0005515; F:protein binding; IEA.
 DR InterPro; IPR004093; Staphylokinase.
 DR InterPro; IPR008124; Streptokinase.
 DR Pfam; PF02821; Staphylokinase; 3.
 DR PRINTS; PR01753; STREPKINASE.
 SQ SEQUENCE 414 AA; 47254 MW; F75B5831B766904 CRC64;

Alignment Scores:
 Pred. No.: 7, 96e-140 Length: 414
 Score: 2129.00 Matches: 408
 Percent Similarity: 99.03% Conservative: 2
 Best Local Similarity: 98.55% Mismatches: 4
 Query Match: 67.93% Indels: 0
 DB: 2 Gaps: 0

US-09-940-235-11 (1-1782) x Q53284_STREBQ (1-414)

OY 538 ATTGCTGACGCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACGCAATTGTTGTT 597
 Db 1 MetAlaGlyProGluThrLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal 20
 OY 558 AGCGTTGCTGTGACTGTTGAGGGGACGATCAAGCACTTATGCTTAAATTTTGAATC 657
 Db 21 SerValAlaGlyThrValGlnGlnLeuThrAsnGlnAspIleSerLeuLysPheGluIle 40
 OY 658 GATCTAACATCAAGCACTGCTGATGAGGAAGCAAGCAAGGCTTAATCCAAATCA 717
 Db 41 AspleuThrSerArgProAlaHisGlyGlyLeuThrGlnGlnLeuSerProLysSer 60
 OY 718 AAACCATTTGCTACTGATGATGCGCGGATGTCAATTAATTTGAGAAAGCTGACTTACTA 777
 Db 61 LysProPheAlaThrAspSerGlyAlaMetProHisLeuGlnLysAlaAspLeu 80
 OY 778 AAGGCTATTCAAGAACATTTGATGCTTAACGTCACAGTAAACGACGACTTCTTGAAGTC 837
 Db 81 LysAlaIleGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGlnVal 100
 OY 838 ATTGATTTTGGCAAGCATGACCACTTACTGATCGAAAGCGCAAGGCTTACTTTGCTGAC 897
 Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyValTyrPheAlaAsp 120
 OY 898 AAAAGATGGTTCGGTAACTTGGCCGACCCCAACGTCGCAAGATTTTGTCTAAGCGGACAT 957
 Db 121 LysAspGlySerValThrLeuProMetGlnProValGlnGlnPheLeuLeuSerGlyHis 140
 OY 958 GTGCGCGTTGACCAATATAAGAAAAACAATCAACAACCAAGCGAAATCTGTTGATGTG 1017
 Db 141 ValArgValArgProTyrLysGlyLysProIleGlnAsnGlnAlaLysSerValAspVal 160
 OY 1018 GAATTAATCTGTAAGTTTACTCCCTTAAACCTGATGACGATTTTGACACGAGTCTCAAA 1077
 Db 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
 OY 1078 GATTAATGCTATTGAAAACACATAGCTATCGGTGACACCATCACTCAAGAAATTACTA 1137
 Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeuLeu 200
 OY 1138 GCTCAAGCAAGCAATTTTAAACAAAACCAACCCGAGGCTATCGATTATGAAAGTGCAC 1197
 Db 201 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGlnArgAsp 220

OY 1198 TCCTCAATCGTCACTGACATGACAAATGACATTTTCCGTACGATTTTACCAATGGATCAAG 1257
 Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnIle 240
 OY 1258 TTACTTACCGGTGTTAAAAATCGGAAACAGCTTATAGCAATCAATAAAAATCTGGTCTG 1317
 Db 241 PheThrTyrHisValLysAsnArgGlnGlnAlaTyrGlnIleAsnLysSerGlyLeu 260
 OY 1318 AATGAAGAAATTAACAACATGACCTGATCTCTGAGAAATTTAGTCCCTTAAAAAGG 1377
 Db 261 AsnGlnGlnIleAsnAsnThrAspLeuIleSerGlnLysTyrValLeuLysLysGly 280
 OY 1378 GAAAGCGGTATGATCCCTTGTATGCGAGTCACTTGAACGTTCAACCATTAATACGTT 1437
 Db 281 GluLysProTyrAspProPheAspArgSerIleLeuLysLeuPheThrIleLysTyrVal 300
 OY 1438 GATGTGATACCAAGCAATTCCTAATAAGTACGACTTTTACAGCTAGCGAAAGTAAAC 1497
 Db 301 AspValAsnThrAsnGlnLeuLeuLysSerGlnGlnLeuLeuThrAlaSerGlnArgAsn 320
 OY 1498 TTAGACTTCAGAGTTTATATACGATCTCTGTATAGGCTTAAGCTTAATCTTACCAATCTC 1557
 Db 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLysTyrAsnAsnLeu 340
 OY 1558 GATGCTTTGGTATTATGACCTATACCTTAAGTGAAGAGTAAATCACAGATGAC 1617
 Db 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlnLysValGlnAspAsnHisAspAsp 360
 OY 1618 ACCAACCGTATCATACCGTTTATATAGGCAAGGACCCGAAGAGCAAGATGCTAGTAT 1677
 Db 361 ThrAsnArgIleIleThrValTyrMetGlyAsnArgProGlnGlnGlnAlaSerTyr 380
 OY 1678 CATTACCTTATGATTAATAAGTGGTATATCCGAAGAAGAAGCAAGATTTCAGCTTACCTG 1737
 Db 381 HisLeuAlaTyrAspLysAspArgTyrThrGlnGlnGlnValLysTyrSerTyrLeu 400
 OY 1738 CGTTATACAGGACACACTATACCTGATACCTTAACCTTAACGACAA 1779
 Db 401 ArgTyrThrGlnThrProIleProAspAsnProAsnAspLys 414

RESULT 3
 ID 06UK57_STREBQ PRELIMINARY; PRT; 436 AA.
 AC 06UK57;
 DT 05-JUL-2004 (TREMBLrel. 27, Created)
 DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
 DE Streptokinase.
 OS Streptococcus equisimilis.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_Taxid=119602;
 RN (1)
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=ATCC 35666;
 RA Costa C.S., Torres F.A.G., Filho S.A.;
 RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY68335; A073571.1; -; Genomic_DNA.
 DR SMR; 06UK57; 59-394.
 DR GO; GO:0016301; F:kinase activity; IEA.
 DR GO; GO:0008243; F:plasmaeogen activator activity; IEA.
 DR GO; GO:0005515; F:protein binding; IEA.
 DR InterPro; IPR004093; Staphylokinase.
 DR InterPro; IPR008124; Streptokinase.
 DR Pfam; PF02821; Staphylokinase; 3.
 DR PRINTS; PR01753; STREPKINASE.
 KW Kinase.
 SQ SEQUENCE 436 AA; 49658 MW; CA2D95F957D8F3C CRC64;

Alignment Scores:
 Pred. No.: 5,5e-139 Length: 436
 Score: 2117.00 Matches: 405

Percent Similarity: 98.55%
 Best Local Similarity: 97.83%
 Query Match: 67.55%
 DB: 2
 Gaps: 0

US-09-940-235-11 (1-1782) x Q6UK57_STR04 (1-436)

```

QY 538 ATTGCTGACCTGATGAGTGGCGGATGTCATTAATCTGTCACCAAGCAATGTTGTT 597
DB 23 IleaaglyProgluThreuleuAspArgProSerValAsnAsnSerGlnLeuVal 42
QY 558 ACCGTTGCTGATCTGTTGAGGAGCAATCAAGACATTAATGTTAAATTTTGAATC 657
DB 43 SerValaIaaglyThrValGluGluThrAnGlnAspIleSerLeuysPhePheGluIle 62
QY 658 GATCTTACATCAAGACCTGCTCAAGAGAAAGACAGACAGGCTTAAGTCCAAATCA 717
DB 63 AspLeuThrSerGlnProAlaHisGlyGlyLysThrGlnGlnGlyLeuSerProLysSer 82
QY 718 AAACCATTTGCTACTGATGATGCGGCGGATGTCATTAATCTGTCACCAAGCTTAATCA 777
DB 83 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGlyValAspLeuLeu 102
QY 778 AAGGCTATTCAGAAACAATTTGATCGCTAAAGTCCACAGATCAAGACATCTTGAAGTC 837
DB 103 LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyrrPheGluVal 122
QY 838 ATTGATTTTGCAGGAGTCAACCATTAATCTGATCAAGAGGAGGCTTAATCTTGTGAC 897
DB 123 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyValValTyrrPheAlaAsp 142
QY 898 AAAGATGCTGGTGAACCTTGGCGGACCCGATCTGTCACCAAGATTTTGTGAAGCAAT 957
DB 143 LysAspGlySerValThrLeuProThrGlnProValGlnGlnLeuLeuLysGlyHis 162
QY 958 GTGCGCGTTAGACCAATATTAAGAAAAACAATCAAAACAAGACGAAATCTGTTGATGTC 1017
DB 163 ValaGValArgProGlyArgGlyLysProValGlnAsnGlnAlaLysSerValaAspVal 182
QY 1018 GAATTAATCTGACGTTTACTCCCTTAAACCTGATGACGATTTGACACGAGTCTCAAA 1077
DB 183 GluTyrrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 202
QY 1078 GATTAATGCTTGTGAACCACTAGTATCGGTGACCAACATCAATCTTCAAGATTAATCA 1137
DB 203 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeuLeu 222
QY 1138 GCTCAAGACCAAGACATTTTAAACAAAACAACCCGAGGCTATPACGATTTATGAACGTGAC 1197
DB 223 AlaGlnAlaGlnSerIleLeuAsnLysThrHisAspGlyTyrrThrIleTyrrGlnAlaArgAsp 242
QY 1198 TCCTCAATCGTCACTCATGACCAATGACATTTTCCGATGATTTTACCAATGATCAAGAG 1257
DB 243 SerSerIleValIleThrHisAspLysAspIlePheArgThrIleLeuProMetAspGlnGln 262
QY 1258 TTTACTTACCGGTTTAAATATGGGAAACAAGCTTATAGATCAATTAATAATTTGGTCTG 1317
DB 263 PheThrTyrrHisValaLysAsnArgGlnGlnAlaTyrrGlnIleAsnLysLysSerGlyLeu 282
QY 1318 AATGAAGAAATTAACAACAAGCTGATCTCTGAAGAAATATTAAGCTTAAAGGAG 1377
DB 283 AsnGlnGlnIleAsnAsnThrAspLeuIleSerGlnLysTyrrTyrrValLeuLysGlyVal 302
QY 1378 GAAAAGCCGATGATCCCTTTGATCGACGATCACTTGAACCTGTTCAACATCAATGATCT 1437
DB 303 GluLysProTyrrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrrVal 322
QY 1438 GATGTCGATCCAAAGAAATTTGCTAAAGATGACAGCTTCTTAAACAGCTACGGAAGTAAAC 1497
DB 323 AspValaAsnThrAsnGlnLeuLeuLysSerGlnGlnLeuThrAlaSerGlnAlaArgAsn 342
QY 1498 TTAGACTTCAGAGATTTATAGATCTCGTGAATTAAGGCTAAATCTTCAACCAATCTTC 1557

```

```

DB 343 LeuAspPheArgAspLeuTyrrAspProArgAspLysAlaLysLeuLeuTyrrAsnAsnLeu 362
QY 1558 GATGCTTTTGGTATTTATGACCTTATACCTTAACTGGAAGAAAGTAAAGATTAACAGATGAC 1617
DB 363 AspAlaPheGlyIleMetAspTyrrThrLeuThrGlyValGlnAspAsnHisAspAsp 382
QY 1618 ACCAAGCCGATCAATTAACCGTTTATATGGGACAGGACCCGAAAGAGAAATGCTAGCTAT 1677
DB 383 ThrAsnArgIleIleThrValIleTyrrMetGlyLysArgProGlnGlnGlnAlaSerTyrr 402
QY 1678 CATTAGCCTTATGATTAAGATCGTTATATCCGAAAGAAACAAGAGATTACAGCTACTG 1737
DB 403 HisLeuAlaTyrrAspLysAspArgTyrrThrGlnGlnGlnAlaValTyrrSerTyrrLeu 422
QY 1738 CGTTTATACAGGACCACTTATACCTGATTAACCTTAACGACAA 1779
DB 423 ArgTyrrThrGlyThrProIleProAspAsnProAspAspLys 436

```

RESULT 4

```

STRP_STRS1
ID STRP_STRS1 STANDARD; PRT; 440 AA.
AC P10519;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Streptokinase G precursor.
GN Name=SkG;
OS Streptococcus sp. (strain 19909).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=69017;
RN [1] NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RP MEDLINE=89160265; PubMed=2922269;
RA Walter F., Siegel M., Maile H.;
RT "Nucleotide sequence of the streptokinase gene from a group-G
RT Streptococcus".
RL Nucleic Acids Res. 17:1262-1262(1989).
CC -!- FUNCTION: This protein is not a protease, but it activates
CC plasminogen by complexing with it. As a potential virulence
CC factor, it is thought to prevent the formation of effective fibrin
CC barriers around the site of infection, thereby contributing to the
CC invasiveness of the cells.

```

CC This Swiss-Prot entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

```

CC -----
CC EMBL, X13400; CAA31766.1; -; Genomic DNA.
CC PIR, S02723; S02723.
CC HSSP, P00779; 1QOR.
CC SMK, P10519; 63-398.
CC InterPro: IPR004093; Staphylokinase.
CC InterPro: IPR008124; Streptokinase.
CC Pfam: PF02821; Staphylokinase; 3.
CC PRINTS; PR01753; STREPKINASE.
CC KM Plasminogen activation; Signal; Virulence.
CC FT SIGNAL 1 26
CC CHAIN 27 440 Streptokinase G.
CC SQ SEQUENCE 440 AA; 50199 MW; 5521F8825FE1B6EA CRC64;

```

Alignment Scores:

```

Pred. No.: 6,46e-139 Length: 440
Score: 2116.00 Matches: 406
Percent Similarity: 98.55% Conservative: 2
Best Local Similarity: 98.07% Mismatches: 6
Query Match: 67.52% Indels: 0
DB: 1 Gaps: 0

```

US-09-940-235-11 (1-1782) x STRP_STRS1 (1-440)

```

OY 538 ATTGCTGGAACCTGAGTGGCTGCTAGAACCTGCTCATCTGTCAACAACAGCAATTGGTTGTT 597
DB 27 TlealaglyProciulntrleuAspaArgProSerValaAsnSergInleuVala 46
OY 598 AGCGTTGCTGTAACCTTTGAGGGGAGCAATCAGACATTAGTCTTAAATTTTGAATC 657
DB 47 SerValaIaglyThrValaIguglyThrAsnGlnAspIleSerleuysPhePheGluIle 66
OY 658 GATCTAACATCAGCACTGCTCATGAGAGAAAGACAGAGCAAGGCTTAAGTCCAAATCA 717
DB 67 AspLeuThrSerArgProAlaHisGlyGlyLeuThrGluGlnGlnleuSerProLysSer 86
OY 718 AAACATTGCTATCTGATAGTGGCGGATGTCATATACTGAGAAAGCTGACTTACTA 777
DB 87 LysLeuPheAlaThrAspSergIylalameProHisLysleuGlnLysalAspLeuLeu 106
OY 778 AAGGCTATTCAGAAACAATTTGATCGCTTAAGTCAAGTACAGCACTTCTTGAAGTC 837
DB 107 LysAlaIleGlnGlnleuInleuIleAlaAsnValHisSerAsnAspPyrPheGluVal 126
OY 838 ATTGATTTTGAAGCGATGCAACATTAATCTGATGAAAGCGCAAGGCTACTTGTGCTGAC 897
DB 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyValIlyrPheAlaAsp 146
OY 898 AAAGATGGTTCGGTAACTTGGCCGACCAACCTGTCAGAAATTTTGTAAAGCGGACAT 957
DB 147 LysAspGlySerValThrleuProIleGlnProValGlnleuLeuLeuLysGlyHis 166
OY 958 GTGCGCGTTAGACCATATTAAGAAAAACCAATACAAACAGCAAGCAATCTGTTGATG 1017
DB 167 ValaGValaArgProLysGlyLysProValGlnAsnGlnAlaLysSerValAspVal 186
OY 1018 GAATPACTGTGACAGTTTCTCCCTTAAACCCGATGAGATTTGACACAGCTTCACAA 1077
DB 187 GluIlyrThrValaGlnPheThrProLeuAsnProAspAspPheArgProAlaLeuLys 206
OY 1078 GATCTAAGCTATTTGAAAACTAGCTATCGGTGACACCATCTCAAGATTAATCTA 1137
DB 207 AspThrLysleuLeuLysThrleuAlaIleGlyAspThrIleThrSergInleuLeu 226
OY 1138 GCTGACGACAAAGCATTTTAAACAAAACCAACCAAGGCTATACGATTAATGAACGTGAC 1197
DB 227 AlaGlnAlaGlnSerIleleuSndySAsnHisProGlyIlyrThrIleIyrgluAsp 246
OY 1198 TCCTCAATGTGACTCTAGACATGACATTTTCCGATCGATTTTACCATGATGACAG 1257
DB 247 SerSerIleValaIlyrHisAspAsnAspIlePheArgThrIleleuProLysSerGln 266
OY 1258 TTTACTTACCGTGTAAAAATGCGGAAACAAGCTTATAGGATCAATAAAAAATCGGTCTG 1317
DB 267 PheThrIyHisValaLysAsnAspGlnGlnAlaIyrgluLeuSndyLysSerGlyLeu 286
OY 1318 AATGAAGAATTAACAACAACATGACCTGATCTGTGAGAAATATTAATCTCTTAAAAAGG 1377
DB 287 AsnGluGluIleAsnAsnThrAspLeuIleSergIulysIyrgluValleuLysGly 306
OY 1378 GAAAAGCGGTATGATCCCTTTGATGCGAGTCACTGAAAATGTTCAACATCAATCGTT 1437
DB 307 GluLysProLysAspProPheAspArgSeriIleLysleuPheThrIleLysIyrgluVal 326
OY 1438 GATGTCGATACCAAGCAATTTGCTTAAAGTGAAGCACTTAACTGACCTTAAACCTTAC 1497
DB 327 AspValaAsnThrAsnGlnleuLeuLysSergIulGlnleuLeuThrAlaSergIulAspAsn 346
OY 1498 TTAGACTTCAGAGATTATAGCATCTGCTGATTAAGGCTTAACTACTTACCAACATCTC 1557
DB 347 LeuAspPheArgAspLeuIyAspProAlaHisGlyLysThrGlnGlnleuSerProLysSer 366
OY 1558 GATGCTTTTGTATTAATGACATATACCTTAATCTGAAAAGTGAAGATATACGATGAC 1617
DB 367 AspAlaPheGlyIleLysAspPyrThrIleuThrGlyLysValaGlnAspAsnHisAspAsp 386
OY 1618 ACCAACCGTATCATTAACGTTTATATATGCGCAAGCGACCCGAAAGAGAAATGCTAGCTAT 1677

DB 387 ThrAsnArgIleIleThrValIyLysArgProGlnGlnAlaSerIyrgly 406
OY 1678 CATTACCTATGATTAAGATCGCTTATACCAAGAAAGCAAGAAAGTTTAACTACTCTG 1737
DB 407 HisLeuAlaIyAspLysAspArgIyThrGlnGlnGlnleuLysGlnValIySeryrLeu 426
OY 1738 CGTTATACAGGAACCACTATCTGATTAACCTTAACTGACGACAA 1779
DB 427 ArgIyThrGlnIlyrThrProIleProAspAsnProAsnLys 440

RESULT 5
OY0X0Y1 STRPY
ID O7X0Y1 STRPY PRELIMINARY; PRT; 440 AA.
AC O7X0Y1;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Streptokinase.
GN Name=eka;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=D488;
RX PubMed=14679231; DOI=10.1128/JB.186.1.110-121.2004;
RA Kalia A., Bessen D.E.;
RT "Natural selection and evolution of streptococcal virulence genes
RT involved in tissue-specific adaptations." ;
RL J. Bacteriol. 186:110-121(2004)
DR EMBL; AY234137; AAP39957.1; -; Genomic_DNA.
DR HSSP; Q53284; 1C4P.
DR SMK; Q7X0Y1; 63-398.
DR GO; GO:0016301; F:Kinase activity; IEA.
DR GO; GO:0008243; F:Plasminogen activator activity; IEA.
DR GO; GO:0005515; F:Protein binding; IEA.
DR InterPro; IPR004093; Staphylokinase.
DR Pfam; PF02821; Staphylokinase.
DR PRINTS; PR01753; STREPKINASE.
KW Kinase.
SQ SEQUENCE 440 AA; 49903 MW; CC4BE69647043BAC CRC64;

Alignment Scores:
Pred. No.: 2.83e-127 Length: 440
Score: 1949.00 Matches: 372
Percent Similarity: 94.20% Conservative: 18
Best Local Similarity: 89.86% Mismatches: 24
Query Match: 62.19% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-11 (1-1782) x O7X0Y1_STRPY (1-440)
OY 538 ATTGCTGGAACCTGAGTGGCTGCTAGAACCTGCTCATCTGTCAACAACAGCAATTGGTTGTT 597
DB 27 TlealaglyProciulntrleuAspaArgProSerValaAsnSergInleuVala 46
OY 598 AGCGTTGCTGTAACCTTTGAGGGGAGCAATCAGACATTAGTCTTAAATTTTGAATC 657
DB 47 SerValaIaglyThrValaIguglyThrAsnGlnAspIleSerleuysPhePheGluIle 66
OY 658 GATCTAACATCAGCACTGCTCATGAGAGAAAGACAGAGCAAGGCTTAAGTCCAAATCA 717
DB 67 AspLeuThrSerArgProAlaHisGlyGlyLeuThrGluGlnGlnleuSerProLysSer 86
OY 718 AAACATTGCTATCTGATAGTGGCGGATGTCATATACTGAGAAAGCTGACTTACTA 777
DB 87 LysLeuPheAlaThrAspSergIylalameProHisLysleuGlnLysalAspLeuLeu 106
OY 838 ATTGATTTTGAAGCGATGCAACATTAATCTGATGAAAGCGCAAGGCTACTTGTGCTGAC 897
DB 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyValIlyrPheAlaAsp 146
OY 898 AAAGATGGTTCGGTAACTTGGCCGACCAACCTGTCAGAAATTTTGTAAAGCGGACAT 957
DB 147 LysAspGlySerValThrleuProIleGlnProValGlnleuLeuLeuLysGlyHis 166
OY 958 GTGCGCGTTAGACCATATTAAGAAAAACCAATACAAACAGCAAGCAATCTGTTGATG 1017
DB 167 ValaGValaArgProLysGlyLysProValGlnAsnGlnAlaLysSerValAspVal 186
OY 1018 GAATPACTGTGACAGTTTCTCCCTTAAACCCGATGAGATTTGACACAGCTTCACAA 1077
DB 187 GluIlyrThrValaGlnPheThrProLeuAsnProAspAspPheArgProAlaLeuLys 206
OY 1078 GATCTAAGCTATTTGAAAACTAGCTATCGGTGACACCATCTCAAGATTAATCTA 1137
DB 207 AspThrLysleuLeuLysThrleuAlaIleGlyAspThrIleThrSergInleuLeu 226
OY 1138 GCTGACGACAAAGCATTTTAAACAAAACCAACCAAGGCTATACGATTAATGAACGTGAC 1197
DB 227 AlaGlnAlaGlnSerIleleuSndySAsnHisProGlyIlyrThrIleIyrgluAsp 246
OY 1198 TCCTCAATGTGACTCTAGACATGACATTTTCCGATCGATTTTACCATGATGACAG 1257
DB 247 SerSerIleValaIlyrHisAspAsnAspIlePheArgThrIleleuProLysSerGln 266
OY 1258 TTTACTTACCGTGTAAAAATGCGGAAACAAGCTTATAGGATCAATAAAAAATCGGTCTG 1317
DB 267 PheThrIyHisValaLysAsnAspGlnGlnAlaIyrgluLeuSndyLysSerGlyLeu 286
OY 1318 AATGAAGAATTAACAACAACATGACCTGATCTGTGAGAAATATTAATCTCTTAAAAAGG 1377
DB 287 AsnGluGluIleAsnAsnThrAspLeuIleSergIulysIyrgluValleuLysGly 306
OY 1378 GAAAAGCGGTATGATCCCTTTGATGCGAGTCACTGAAAATGTTCAACATCAATCGTT 1437
DB 307 GluLysProLysAspProPheAspArgSeriIleLysleuPheThrIleLysIyrgluVal 326
OY 1438 GATGTCGATACCAAGCAATTTGCTTAAAGTGAAGCACTTAACTGACCTTAAACCTTAC 1497
DB 327 AspValaAsnThrAsnGlnleuLeuLysSergIulGlnleuLeuThrAlaSergIulAspAsn 346
OY 1498 TTAGACTTCAGAGATTATAGCATCTGCTGATTAAGGCTTAACTACTTACCAACATCTC 1557
DB 347 LeuAspPheArgAspLeuIyAspProAlaHisGlyLysThrGlnGlnleuSerProLysSer 366
OY 1558 GATGCTTTTGTATTAATGACATATACCTTAATCTGAAAAGTGAAGATATACGATGAC 1617
DB 367 AspAlaPheGlyIleLysAspPyrThrIleuThrGlyLysValaGlnAspAsnHisAspAsp 386
OY 1618 ACCAACCGTATCATTAACGTTTATATATGCGCAAGCGACCCGAAAGAGAAATGCTAGCTAT 1677

```



```

Db      107  LysAlaIleGlnLysGlnLeuIleAlaAsnValHisSerAsnAspLysTyrPheGluVal 126
Qy      838  ATTGATTTTGAAGCGATGCAACCATTTATGATGCAAGGCGAGGCTTATGCTTGAC 897
Db      127  IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlnValTyrPheAlaAsp 146
Qy      898  AAAAGATGCTTCGGTAACTTGGCCGACCCAACTGTCCAAAGATTTTGGTAAACGACAT 957
Db      147  LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlnHis 166
Qy      958  GTGGCGGCTTGAACCATTTAAAGAAAAACCATTCACAAACGACGAAATCTGTTGATGCG 1017
Db      167  ValArgValArgProTyrLysGlnLysProValGlnAsnGlnAlaLysSerValAspVal 186
Qy      1018  GAAATATACGTGACAGTTTACCTCCCTTAAACCCGATGACGATTTTCAGACACGCTCAAA 1077
Db      187  LysTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlnLysLeuLys 206
Qy      1078  GATTAAGCTATTTGAAAAACATAGCTATCGGTGACACCATCATCTCAAGAAATTACTTA 1137
Db      207  AspThrLysLeuLeuLysThrLeuAlaIleGlnLysPheThrIleThrSerGlnGluLeu 226
Qy      1138  GCTCAAGCAACAAAGCATTTTAAACAAAACCAAGCCGCTATACGATTTATGAACTGAC 1197
Db      227  AlaGlnAlaGlnSerIleLeuLysLysThrHisProGlnTyrThrIleTyrGlnLysArg 246
Qy      1198  TCCCTAATCGTCACTCATGACATGACATTTTCCGATACGATTTTACCATGATGACAAAG 1257
Db      247  SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGln 266
Qy      1258  TTTACTTACCGTGTAAAAATCGGGAACAGCTTTATAGGATCAATAAAAAATCTGCTG 1317
Db      267  PheThrTyrArgValLysAspArgGlnGlnAlaTyrGlnIleAsnLysSerGlnLys 286
Qy      1318  AATGAGAAATTAACAACAACCTGACCTGATCTGTGAGAAAATATTATCGTCTTAAAAAGG 1377
Db      287  AsnGlnGlnIleAsnAsnThrAspLeuIleSerGlnLysTyrIleLeuLysLysGln 306
Qy      1378  GAAAGCCGATATGATCCCTTATGATCGGATGCTTGAACCTGTTCACCATCAATACGTT 1437
Db      307  GlnSerProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326
Qy      1438  GATGTCGATACCAAGCAATTTGCTTAAAAAGTGAAGAGCTCTTAAACACTGCGCAAGTAC 1497
Db      327  AspValAsnThrAsnGlnLeuLeuLysSerGlnGlnLeuLeuThrAlaSerGlnLysGln 346
Qy      1498  TTAAGCTTCAGAGATTATATGATCTCTCGTGAATAGGCTTAACTACTTACAACTATCTC 1557
Db      347  LeuAspPheArgAspLeuLysArgProArgAspLysAlaLysLeuLeuLysAsnLeu 366
Qy      1558  GATGCTTTTGGTATTATGACATATACCTTAACTGGAAGAAAGTGAAGATATATCAAGTAC 1617
Db      367  AspAlaPheAspIleMetAspTyrThrIleLeuThrGlnLysValGlnAspAsnHisAspLys 386
Qy      1618  ACCAAGCCGATATCAACCGTTATATAGGGCAAGCGACCGGACGAGAGAGATGCTACTCT 1677
Db      387  AsnAsnAspGlnLeuAlaThrValTyrMetGlnLysArgProLysGlnAlaLysGlnSerTyr 406
Qy      1678  CATTAGCTTATGATTAAGATCGTTATACCGAAGAGAAGACGATTTACAGTCACTGCG 1737
Db      407  HisLeuAlaTyrAspLysAspProTyrThrGlnGlnGlnLysAlaTyrSerTyrLeu 426
Qy      1738  CGTTATACAGGAGCACTTACCTGATTAACCTTAAACGACAA 1779
Db      427  ArgAspThrGlnThrProIleProAspAsnProLysAspLys 440

```

```

DE      Streptokinase A precursor.
GN      Name=ska; OrderedLocNames=SPY1979;
OS      Streptococcus pyogenes.
OC      Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC      Streptococcus.
OX      NCBI_TaxID=1314;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RC      STRAIN=SF130/13 / Serotype M1;
RX      MEDLINE=89160264; PubMed=2646590;
RA      Walter F., Siegel M., Walke H.;
RT      "Nucleotide sequence of the streptokinase gene from a Streptococcus
RL      pyogenes type 1 strain."
RL      Nucleic Acids Res. 17:1261-1261(1989).
RN      [2]
RP      NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC      STRAIN=SF370 / ATCC 700294 / Serotype M1;
RX      MEDLINE=1192684; PubMed=11296296; DOI=10.1073/pnas.071559398;
RA      Ferracci J.J., Mcshan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA      Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA      Qian Y., Jia H.G., Najjar P.Z., Ren Q., Zhu H., Song L., White J.,
RA      Yuan X., Clifton S.W., Roe B.A., McLaughlin R.E.;
RT      "Complete genome sequence of an M1 strain of Streptococcus pyogenes."
RC      Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
CC      -1- FUNCTION: This protein is not a protease, but it activates
CC      plasminogen by complexing with it. As a potential virulence
CC      factor, it is thought to prevent the formation of effective fibrin
CC      barriers around the site of infection, thereby contributing to the
CC      invasiveness of the cells.
CC      -----
CC      This Swiss-Prot entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use as long as its content is in no way modified and this statement is not
CC      removed.
CC      -----
DR      EMBL: X13399; CAA31765.1; -; Genomic DNA.
DR      EMBL: AB006620; AAK34665.1; -; Genomic DNA.
DR      PIR: S02724; S02724.
DR      HSSP: Q53284; 1CAP.
DR      SMR: P10520; 63-398.
DR      InterPro: IPRO04093; Staphylokinase.
DR      InterPro: IPRO08124; Streptokinase.
DR      Pfam: PF02821; Staphylokinase; 3.
DR      PRINTS: PR01753; STREPKINASE.
KW      Complete proteome; Plasminogen activation; Signal; Virulence.
FT      SIGNAL 1 26
FT      CHAIN 27 440
FT      CONFLICT 163 163 L -> V (in Ref. 1).
FT      CONFLICT 345 345 R -> G (in Ref. 1).
FT      CONFLICT 373 373 D -> N (in Ref. 1).
FT      CONFLICT 428 428 D -> Y (in Ref. 1).
FT      CONFLICT 438 438 K -> N (in Ref. 1).
SQ      SEQUENCE 440 AA; 49924 MW; D6227EF040B758DB CRC64;

Alignment Scores:
Pred. No.: 1,94e-126 Length: 440
Score: 1937.00 Matches: 367
Percent Similarity: 93.48% Conservative: 20
Best Local Similarity: 88.65% Mismatches: 27
Query Match: 61.81% Indels: 0
DB: 1 Gaps: 0

US-09-940-235-11 (1-1782) x STRP_STRPY (1-440)
Qy      538  ATTGCTGAGCGCTGAGTGGCTGACCGCTCATCTGTCACAAAGCGCAATTTGTTGTT 597
Db      27  IleAlaGlnTyrGlnTyrLeuProAspArgProIleAsnAsnSerGlnLeuValVal 46
Qy      598  AGCGTTGCTGACTGCTTGAAGGAGCAATCAAGACATTAAGCTTAATTTTGAATC 657
Db      47  SerMetAlaGlnIleValGlnGlnThrAspLysLysValPheIleAsnPhePheGlnIle 66

```



```

QY 658 GATCTAATCATCAGCACTGCTCATGAGAGAAAGACAGCAAGGCTTAAGTCCAAATCA 717
DB 67 AspleuThrserrInProAlahIsGlyGlySerThrgInGInGlyLeuSerProLySer 86
QY 718 AAACCAATTTGCTAGTAGTGAGGCGGATGTCATATAACTTGAAGAAAGCTTAATCTA 777
DB 87 LysProPheAlaThrAspAsnGlyAlaMetProHisLysLeuGluValAspLeuLeu 106
QY 778 AAGCTATTTCAGAAACATTTGATCGCTAACGTCACAGTAACGACGACTTACTTTGAGGTC 837
DB 107 LysAlaIleGlnLysGlnLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal 126
QY 838 ATTTGATTTTGAAGCGATGCAACCATTAAGTGAAGCGGCAAGGTCTTACTTTGCTGAC 897
DB 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyValTyrPheAlaAsp 146
QY 898 AAAGATGTTGGTGAACCTTGGCCGACCACTGTCCAAAGATTTTGTCTAAGCGGACAT 957
DB 147 LysAspGlySerValThrLeuProThrGlnProValGlnGlnPheLeuLeuLysGlyHis 166
QY 958 GTGCGCGTTAGACCATATAAAGAAAAACAATCAAAACCAAGGAAATGTTGATGTG 1017
DB 167 ValArgValArgProTyrLysGluLysProValGlnAsnGlnAlaLysSerValAspVal 186
QY 1018 GAATATACCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTCAGACCAAGTCTCAAA 1077
DB 187 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 206
QY 1078 GATCTAAGCTATTGAAAAACAATGATGCTATGGTGACACCATCATCTCAAGATTAATA 1137
DB 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 226
QY 1138 GCTCAAGCAACAAGCATTTTAAACAACCAACCCAGGCTATTCGATTAATGAAACGAGAC 1197
DB 227 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp 246
QY 1198 TCTCATGCTGTCATGATGACATGATCATTTTCCGTACGATTTTCAAGATGATCAAGAG 1257
DB 247 SerSerIleValIleHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
QY 1258 TTTTACTTACCGGTGTTAAAAATCGGAAACAAGCTTATGATCAATAAAAATCTGCTGTG 1317
DB 267 PheThrTyrHisValLysAsnArgGlnGlnAlaTyrGluIleAsnProLysThrGlyIle 286
QY 1318 AATGAAGAATAAACAACACTGACCTGATCTGAGAAATATTAACGCTTAAAAAGGG 1377
DB 287 LysGluLysThrAsnAsnThrAspLeuValSerGluLysTyrTyrValLeuLysGlnGly 306
QY 1378 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCACCATCAATAGCT 1437
DB 307 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326
QY 1438 GATGTGATACCAACGAATGCTAAAAAGTAGACAGCTCTTAAACAGCTACGGAACGTAC 1497
DB 327 AspValAsnThrAsnGlnLeuLeuLysSerGlnGlnLeuLeuThrAlaSerGluArgAsn 346
QY 1498 TTAAGACTTACAGATTTATACATCTCTGATGAAGGCTTAATCACTCTTCAACAATCTC 1557
DB 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
QY 1558 GATGCTTTGGTATTATGACTTACCTTAACCTGAAAGATAGAGATATTCAGATGAC 1617
DB 367 AspAlaAspAspIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspLys 386
QY 1618 ACCAACCCTATACATAACGTTTATATGCGCAAGCGACCCGACGAGAAATGCTGATAT 1677
DB 387 AsnAsnArgValValThrValTyrMetGlyLysArgProLysGlyAlaLysGlySerTyr 406
QY 1678 CATTTAGCTTATGATTAAGATCGTTATACCGAAGAAAGACGAGAAATTTACACTACCTG 1737
DB 407 HisLeuAlaIleTyrAspLysAspLeuTyrThrGlnGlnGluArgLysAlaTyrSerTyrLeu 426
QY 1738 CGTTATACGAGGACACTTACTGATTAACCTTAACGACGAAA 1779

```

```

DB 427 ArgAspThrGlyThrProIleProAspAsnProLysAspLys 440
RESULT 7
ID Q53ZK6_STRPY PRELIMINARY; PRT; 440 AA.
AC Q53ZK6;
DT 13-SEP-2005 (TREMblrel. 31, Created)
DT 13-SEP-2005 (TREMblrel. 31, Last sequence update)
DE 13-SEP-2005 (TREMblrel. 31, Last annotation update)
DE Streptokinase.
GN Name=ska;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RX [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=86-779;
RX PubMed=14679231; DOI=10.1128/JB.186.1.110-121.2004;
RA Kalia A., Bessen D.E.;
RT "Natural selection and evolution of streptococcal virulence genes
involved in tissue-specific adaptations."
RL J. Bacteriol. 186:110-121(2004).
DR EMBL; AY234128; AAP39948.1; -; Genomic_DNA.
KW Kinase.
SQ SEQUENCE 440 AA; 49924 MW; D6227BF040B758DB CRC64;

Alignment Scores:
Pred. No.: 1,94e-126 Length: 440
Score: 1937.00 Matches: 367
Percent Similarity: 93.48% Conservative: 20
Best Local Similarity: 88.65% Mismatches: 27
Query Match: 61.81% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-11 (1-1782) x Q53ZK6_STRPY (1-440)
QY 538 ATTGCTGACCTGATGAGTGGCTGCTGACCGTCCATCTGTCAACAACGCGAATGGTGT 597
DB 27 IleAlaGlyTyrGlyTyrPLeuProAspArgProProIleAsnAsnSerGlnLeuVal 46
QY 598 AGCTTCTGCTGACTGTTGAGGGGACGAATCAAGACATTAATTTTGAATC 657
DB 47 SerMetAlaGlyIleValGlnGlyThrAspLysValPheIleAsnPheGluIle 66
QY 658 GATTAACATCAACGACCTGCTCATGAGAGAAAGACAGAGCAAGGCTTAAGTCCAAATCA 717
DB 67 AspleuThrserrInProAlahIsGlyGlySerThrgInGInGlyLeuSerProLySer 86
QY 718 AAACCAATTTGCTAGTAGTGAGGCGGATGTCATATAACTTGAAGAAAGCTTAATCTA 777
DB 87 LysProPheAlaThrAspAsnGlyAlaMetProHisLysLeuGluValAspLeuLeu 106
QY 778 AAGCTATTTCAGAAACATTTGATCGCTAACGTCACAGTAACGACGACTTACTTTGAGGTC 837
DB 107 LysAlaIleGlnLysGlnLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal 126
QY 838 ATTTGATTTTGAAGCGATGCAACCATTAAGTGAAGCGGCAAGGTCTTACTTTGCTGAC 897
DB 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyValTyrPheAlaAsp 146
QY 898 AAAGATGTTGGTGAACCTTGGCCGACCACTGTCCAAAGATTTTGTCTAAGCGGACAT 957
DB 147 LysAspGlySerValThrLeuProThrGlnProValGlnGlnPheLeuLeuLysGlyHis 166
QY 958 GTGCGCGTTAGACCATATAAAGAAAAACAATCAAAACCAAGGAAATGTTGATGTG 1017
DB 167 ValArgValArgProTyrLysGluLysProValGlnAsnGlnAlaLysSerValAspVal 186
QY 1018 GAATATACCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTCAGACCAAGTCTCAAA 1077
DB 187 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 206

```

```

QY 1078 GATACAAAGCTATGAAACACTAGCTATCGTGACACCATCATCTCAAGAAATTCTA 1137
DB 207 AapThrvLeuLeuLeuYsthrLeuAlaileGlyAspThrIleThrserrGlnGluLeu 226
QY 1138 GCTCAAGCAACAAGCTTTTAAACAACACCCAGGCTATACGATTATTAAGACGTGAC 1197
DB 227 AlaGlnAlaGlnSerIleLeuLeuYsthrHisProGlyTyrThrIleTyrGlnValGAsp 246
QY 1198 TCCTCAATGTCATCATGATGACATGATCTTCCTGACATTTTACCATGATGATCAAGAG 1257
DB 247 SerSerIleValThrHisAspAsnSerIlePheArgThrIleLeuProMetAspGlnGlu 266
QY 1258 TTACTCTACCGTGTAAAAATCGGGAACAGCTTATAGATCATATAAAAAATCGGTGCT 1317
DB 267 PheThrTyrHisValIleValAsnAsnGlnGlnAlaTyrGlnIleAsnProLysThrGlyIle 286
QY 1318 AATGAGAATAAACAACAACATGACCTGATCTGTGAGAAATATTACGTCCTTAAAAAGG 1377
DB 287 LysGlnIlyThrAsnAsnThrAspLeuValSerGlnLysTyrTyrValLeuLysGlnGly 306
QY 1378 GAAAGCCGATGATGATCCCTTTGATGCGAGTCACTGAAACTGTTCAACATCAATCGTT 1437
DB 307 GlnLysProLysAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326
QY 1438 GATGCGATACCAAGCAATTTGCTAAAAAGTGAAGCAGCTCTTAAACAGTTCGCGAAGCTAAC 1497
DB 327 AspValAsnThrAsnIleuLeuLeuYsterrGlnGlnLeuThrIleAspGlnValArgAsn 346
QY 1498 TTAGACTTCAGAGATTTATACGATCCTCGTGAATAGGCTAAACTACTATCAACAATCTC 1557
DB 347 LeuAspPheArgAspLeuTyrAspProArgAspIlyAlaLysLeuLeuTyrAsnAsnLeu 366
QY 1558 GATGCTTTTGATTTATGACATATACCTTAACCTGAAAAAGTGAAGATTAATCAAGATGAC 1617
DB 367 AspAlaPheAspIleMetAspTyrThrLeuThrGlnIlyValGlnLysAsnHisAspLys 386
QY 1618 ACCAACCAGTATCAACCGTTATATAGGCGAAGCGAAGCGAAGAGAAATGCTACTACTAT 1677
DB 387 AsnAsnAlaGlnValThrValTyrMetGlyLysArgProLysGlyAlaLysGlySerTyr 406
QY 1678 CATTAGCCTTATGATTAAGATCGTTATACCGAAGAAAGAAAGCTTATACGCTACCTG 1737
DB 407 HisLeuAlaIlyAspLysAspLeuTyrThrGlnGlnGlnIlyAspAlaTyrSerTyrLeu 426
QY 1738 CGTTATACAGGAGACCTTATACCTGATAACCTTAACGACAAA 1779
DB 427 ArgAspThrGlyThrProIleProAspAsnProLysAspLys 440

RESULT 8
Q8K5R8_STRP3 PRELIMINARY; PRT; 440 AA.
ID Q8K5R8_STRP3 PRELIMINARY; PRT; 440 AA.
AC Q8K5R8; Q79W73;
RC STRAIN-MGAS315 / Serotype M3;
RX MEDLINE=22133608; PubMed=12122206; DOI=10.1073/pnas.152298499;
RA Beres S.B., Sylva G.L., Barbican K.D., Lei B., Hoff J.S., Parkins L.D.,
Mammarella N.D., Liu M.-Y., Smoot J.C., Porcilla S.F., Parkins L.D.,
Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
RA Schlievert P.M., Musser J.M.;
RT "Genome sequence of a serotype M3 strain of group A Streptococcus:
RT phage-encoded toxins, the high-virulence phenotype, and clone
RT emergence.";
```

```

RL Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SSI-1 / Serotype M3;
RX MEDLINE=22683278; PubMed=12799345; DOI=10.1101/gr.1096703;
RA Nakagawa I., Kurokawa K., Yamashita A., Nakata M., Tomiyasu Y.,
RA Okahashi N., Kawabata S., Yamazaki K., Shiba T., Yasunaga T.,
RA Hayashi H., Hattori M., Hamada S.;
RT "Genome sequence of an M3 strain of Streptococcus pyogenes reveals a
RT large-scale genomic rearrangement in invasive strains and new insights
RT into phage evolution.";
RL Genome Res. 13:1042-1055 (2003).
DR EMBL; AE014169; AA080305.1; -; Genomic DNA.
DR EMBL; BA000034; BAC64795.1; -; Genomic DNA.
DR HSSP; P00779; 114D.
DR SMK; Q8K5R8; 63-398.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008243; F:plasma membrane activator activity; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR InterPro; IPR004093; Staphylokinase.
DR InterPro; IPR008124; Streptokinase.
DR Pfam; PF02821; Staphylokinase; 3.
DR PRINTS; PR01753; STREPKINASE.
KW Complete proteome; Kinase.
SQ SEQUENCE 440 AA; 49911 MW; 7CCE44F402EE3975 CRC64;

Alignment Scores:
Pred. No.: 1,94e-126 Length: 440
Score: 1937.00 Matches: 372
Percent Similarity: 93.72% Conservative: 16
Best Local Similarity: 89.86% Mismatches: 26
Query Match: 61.81% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-11 (1-1782) x Q8K5R8_STRP3 (1-440)
QY 538 ATTGCTGAGCCTGAGGCGCTGAGACCGCTCATCTGTCAACAACGCAATGGTGT 597
DB 27 IleAlaGlyTyrGlyTyrLeuProAspArgProProValAsnAsnSerGlnLeuVal 46
QY 598 AGCGTTCGTGATCTGTGAGGAGGAGCAACAACATTAAGTCTAAATTTTGAATC 657
DB 47 SerMetAlaGlyIleValGlnGlnIlyThrAspLysValPheIleAsnPheGlnIle 66
QY 658 GATCTAACATCAAGCAGCTGCTCATGAGAGAAAGACAGAGAGGCTTAAGTCAAAATCA 717
DB 67 AspLeuThrSerGlnIleAlaHisGlyGlyLysTyrThrGlnGlnIlyLeuSerProLysSer 86
QY 718 AAACATTTGCTACATGATGCGGCGGATGTCATTAATTTGAGAAAGCTTACTTA 777
DB 87 LysProPheAlaThrAspAsnGlyAlaMetProHisLysLeuGlnLysAlaAspLeu 106
QY 778 AAGGCTATTCAAGAAATGATGCTTAACGTCACAGTAAAGCAGCTACTTTGAGGTC 837
DB 107 LysAlaIleGlnIlySerGlnLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGlnVal 126
QY 838 ATTGATTTTGAAGGAGATGACCATTAATGATGAAAGGCAAGGCTTACTTTGCTGAC 897
DB 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146
QY 898 AAAGATGCTTCGTTAACCCTGCGAAGCAACCTGTCACAAATTTTGTAAACGACAT 957
DB 147 LysAspGlySerValThrLeuProIleGlnProValGlnIlyLeuLeuSerGlyHis 166
QY 958 GTTCGGGTTAGACCATTAATAAGAAACCAATCAAAACCAAGCAAGAAATGTTGATGTG 1017
DB 167 ValArgValArgProTyrIlySerGlnLysProValGlnAsnGlnAlaLysSerValAspVal 186
QY 1018 GAATATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTCAAGAGGCTTCAAA 1077
DB 187 LysTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 206
QY 1078 GATCTAAGCTATTGAAAAACAAGTACGATGCGGAGACCATCATCTCAAGAAATTCTA 1137
```

```

Db      207  AsphThrvLsleuLysThrIleuAlaIleGlyAspThrIleThrsErgIngluLeu 226
QY      1138  GCTCAAGCAAAAGCATTTTAAACAAAACCCCGGCGTATCGATTTATGAACGTGAC 1197
Db      227  AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp 246
QY      1198  TCCTCAATCGTCACTCATGACATGACATTTTCCTGATGATTTTACCAATGATCAAGAG 1257
Db      247  SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
QY      1258  TTTACTTACCGGTGTTAAATGGGAAACAAGCTTATGATCAATATAAAATCTGGCTG 1317
Db      267  PheThrTyrArgValLysAspArgGlnGlnAlaTyrGlyIleAsnLysLysSerGlyLeu 286
QY      1318  AATGAAGAAATTAACAACATGACCTGATCTCTGAGAAATATATACCTTAAATAAGG 1377
Db      287  AsnGlnGlnLysAsnAsnThrAspLeuLysSerGlnLysTyrThrIleLeuLysGly 306
QY      1378  GAAAAGCCGATGATCCCTTGTATGCGACATGCTTGAACCTGTTCAACATACGTT 1437
Db      307  GluSerProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326
QY      1438  GATGTCGATACCAACGAATGCTTAAAGTGAAGCAGCTTTTACAGCTTACGAACTGAC 1497
Db      327  AspValAsnThrAsnGlnLeuLeuLysSerGlnLysGlnLeuThrAlaSerGluArgAsn 346
QY      1498  TTAGCTTCAGAGATTTATACATCTCTGATGATGAGCTTAACTCTTCAACATCTC 1557
Db      347  LeuAspPheArgAspLeuTyrAspProCysAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
QY      1558  GATGCTTTGGTATATGATGACATATACCTTAACTGGAAGATGAGATATACGATGAC 1617
Db      367  AspAlaPheAspIleMetAspPyrThrLeuThrGlyLysValGlnAspAsnHisAspLys 386
QY      1618  ACCAACCGTATCATACCGTTTATATGGGCAAGCGAACCGAAGAGAGATGCTGAT 1677
Db      387  AsnAsnArgIleValThrValTyrMetGlyLysArgProLysGlyAlaLysGlySerTyr 406
QY      1678  CATTTAGCTTATGATTAAGATGTTATATCCGAAAGAAACGAGATTTACAGCTTAC 1737
Db      407  HisLeuAlaTyrAspLysAspLeuTyrThrGlnGlnGluArgLysAlaTyrSerTyrLeu 426
QY      1738  CGTTATACAGGAGACCTTATACCTGATACCTTACCTTACGACAA 1779
Db      427  ArgAspThrGlnThrProIleProAspAsnProLysAspLys 440

RESULT 9
Q7X0Y2_STRPY PRELIMINARY; PRT: 440 AA.
ID  Q7X0Y2_STRPY PRELIMINARY; PRT: 440 AA.
AC  Q7X0Y2_2003 (TrEMBLrel. 25, Created)
DT  01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT  01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE  Streptokinase.
GN  Name=eka;
OS  Streptococcus pyogenes.
OC  Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC  Streptococcus.
OX  NCBI_TaxID=1314;
RN  [1]
RP  NUCLEOTIDE SEQUENCE.
RC  STRAIN=D306;
RC  PubMed=14679231; DOI=10.1128/JB.186.1.110-121.2004;
RA  Kalita A., Bessen D.E.;
RT  "Natural selection and evolution of streptococcal virulence genes
RT  involved in tissue-specific adaptations.";
RL  J. Bacteriol. 186:110-121(2004).
DR  EMBL; AY234136; AAP39956.1; -; Genomic_DNA.
DR  HSBP; O53284; 1C4P.
DR  SMR; O7X0Y2; 63-398.
DR  GO; GO:0016301; F:kinase activity; IEA.
DR  GO; GO:0008243; F:plasminogen activator activity; IEA.

```

```

DR  GO; GO:0005515; F:protein binding; IEA.
DR  InterPro; IPR004093; Staphylokinase.
DR  InterPro; IPR008124; Streptokinase.
DR  Pfam; PF02821; Staphylokinase; 3.
DR  PRINTS; PR01753; STREPKINASE.
SQ  KINASE.
SQ  SEQUENCE 440 AA; 5018 MW; 4CB42AC2A9062C2E CRC64;

Alignment Scores:
Pred. No.: 5,96e-126 Length: 440
Score: 1930.00 Matches: 368
Percent Similarity: 93.72% Conservative: 20
Best Local Similarity: 88.89% Mismatches: 26
Query Match: 61.58% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-11 (1-1782) x Q7X0Y2_STRPY (1-440)
QY      538  ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAAACAGCAATGGTGT 597
Db      27  IleAlaGlyTyrGlyTyrPleuProAspArgProIleAsnAsnSerGlnLeuVal 46
QY      598  AGCGTTCTGCTGATCTGTGAGGGGACGATCAACATTAAGCTTAAATTTTGAATC 657
Db      47  SerMetAlaGlyIleValGlnGlyThrAspLysValPheIleAsnPheGluIle 66
QY      658  GATCTTACATACGACCTGCTCATGAGGAAAGACAGCAAGGCTTAACTGCTCAAAATCA 717
Db      67  AspLeuThrSerGlnProAlaHisGlyLysThrGlnGlnLysLeuSerProLysSer 86
QY      718  AAACATTTGCTACTGATGAGGGCGGATGTCATATACTTGAAGAGCTTACTTA 777
Db      87  LysProPheAlaThrAspAsnSerAlaMetProHisLysLeuGlnLysAlaAspLeuLeu 106
QY      778  AAGGCTATTCAAGAAACATTTGATCGCTACGCTCCAGATTAACGACATCTTGAAGT 837
Db      107  LysAlaIleGlnLysGlnLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal 126
QY      838  ATTGATTTTGCAAGCGATGCAACCATTTACTGATGAAACGGCAAGTCTACTTGGCTGAC 897
Db      127  IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146
QY      898  AAAGATGCTTCGGTAACTTCGCGACCGGACCACTGCTCAAGATTTTGGTAAAGGACAT 957
Db      147  LysAspLysSerValThrLeuProThrGlnProValGlnGlnPheLeuThrGlyHis 166
QY      958  GTGCGCTTACGACATTAAGAAACCAATACCAACCAAGCAAGCAATCTGTTGAAGTG 1017
Db      167  ValArgValArgProTyrLysGlnLysProValGlnAsnGlnAlaLysSerValAspVal 186
QY      1018  GAATATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTCAAGCAAGTCTCAAA 1077
Db      187  GluTyrThrValGlnPheThrProLeuAsnProAspAspPheLysProValLeuLys 206
QY      1078  GATATTAAGCTATTTGAAACATCATGCTTGGTGACACCATCATCATCTCAAGATTA 1137
Db      207  AspThrLysLeuLeuLysLysLeuAlaIleGlyAspThrValThrSerGlnGlnLeuLeu 226
QY      1138  GCTCAAGCAAAAGCATTTTAAACAAAACCAACCCGCGCTATACGATTTTAAAGCTGAC 1197
Db      227  AlaGlnAlaGlnSerIleLeuAsnGlnLysSerHisProAspPyrThrIleTyrGluArgAsp 246
QY      1198  TCCTCAATCGTCACTCATGACATGACATTTTCCTGATGATTTTACCAATGATCAAGAG 1257
Db      247  SerSerIleValThrHisAspLysAspIlePheArgThrIleLeuProMetAspGlnGlu 266
QY      1258  TTTACTTACCGGTGTTAAATGGGAAACAAGCTTATGATCAATATAAAATCTGGCTG 1317
Db      267  PheThrTyrArgValLysAsnAspArgGlnGlnAlaTyrArgIleAsnLysLysSerGlyLeu 286
QY      1318  AATGAAGAAATTAACAACATGACCTGATCTCTGAGAAATATATACCTTAAATAAGG 1377
Db      287  AsnGlnGlnLysAsnAsnThrAspLeuLysSerGlnLysTyrThrIleLeuLysGlyArg 306

```

QY 1378 GAAAACCGTATGATCCCTTGTATGCGAGTCACTTGAAACGTTCACATCAATAGTT 1437
 |||||
 Db 307 GILVSPFOTYRASPProtheAspArgSerHisLeuLysLeuPheThrIleLysIYrVal 326
 QY 1438 GATGTCGATACCAACGAATTTGCTAAAAGTAGACAGCTTTAAACAGCTACGAAAGCTAAC 1497
 |||||
 Db 327 AspValAsnThrAsnGlnLeuLeuLysSerGlnGlnLeuLeuThrAlaSerGlnArgAsn 346
 QY 1498 TTAGACTTCAGAGATTATACGATCCTCGATAGAGCTAACTACTCTCAACAATCTC 1557
 |||||
 Db 347 LeuAspPheArgAspLeuLysIYrAspProAlaGAspLysAlaLysLeuLeuLysIYrAsnAsnLeu 366
 QY 1558 GATGCTTTGTATTTATGACTATACCTTACTACGAAAGAGATAGAGATATACGATAC 1617
 |||||
 Db 367 AspAlaPheAspLleMetAspIYrThrLeuThrGlyLysValGlnAspAsnHisAspLys 386
 QY 1618 ACCAACCGTATCATTAACCGTTATATAGGCAACCGAACCGCAAGAGAGAAATGCTAGCTAT 1677
 |||||
 Db 387 AsnAsnArgValAlaThrValIYrMetGlyLysArgProLysGlyAlaLysGlySerIYr 406
 QY 1678 CATTTAGCTTATGATTAAGATCGTTATACCGAAGAAAGAGAAAGATTACAGTACTG 1737
 |||||
 Db 407 HisLeuAlaIYrAspLysAspLeuLysIYrThrGlnGlnLysArgLysAlaIYrSerIYrLeu 426
 QY 1738 CGTTATACAGGAGACACTTATACCTGATTAACCTTAAGCAAA 1779
 |||||
 Db 427 ArgAspThrGlyThrProIleProAspAsnProLysAspLys 440

RESULT 10

Q7X0Y8_STRPY PRELIMINARY; PRT: 440 AA.

ID Q7X0Y8_STRPY AC Q7X0Y8-
 AC Q7X0Y8-
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Streptokinase.
 GN Name=ska;
 OS Streptococcus pyogenes.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1314;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=88-019;
 RX PubMed=14679231; DOI=10.1128/JB.186.1.110-121.2004;
 RA Kalita A., Bessen D.E.;
 RT "Natural selection and evolution of streptococcal virulence genes
 involved in tissue-specific adaptations.";
 RL J. Bacteriol. 186:110-121(2004)
 DR EMBL, AY234129; AAP3949.1; -; Genomic_DNA.
 DR HSBP; Q53284; 1C4P.
 DR SMR; Q7X0Y8; 63-398.
 DR GO; GO:0016301; P:kinase activity; IEA.
 DR GO; GO:0008243; P:plasmaogen activator activity; IEA.
 DR GO; GO:0005515; P:protein binding; IEA.
 DR InterPro; IPR004093; S:staphylokinase.
 DR InterPro; IPR008124; S:streptokinase.
 DR Pfam; PF02821; S:staphylokinase; 3.
 DR PRINTS; PR01753; STRPKINASE.
 KW Kinase.
 SQ SEQUENCE 440 AA; 49885 MW; 69DB44F4026B3975 CRC64;

Alignment Scores:

Pred. No.: 7e-126 Length: 440
 Score: 1929.00 Matches: 371
 Percent Similarity: 93.48% Conservative: 16
 Best Local Similarity: 89.61% Mismatches: 27
 Query Match: 61.53% Indels: 0
 Gaps: 0

US-09-940-235-11 (1-1782) x Q7X0Y8_STRPY (1-440)

QY 538 ATTGCTGACCTGATGCTGCTAGACCGTCCATCTGTCAACAAGCAATGTTGTT 597
 |||||
 Db 27 IleHisGlyIYrGlyIYrPleuProAspArgProIleValAsnAsnSerGlnLeuVal 46
 QY 598 AGCTTGCTGCTACTGTTGAGGGAGCAATCAAGACTTACTCTTAATTTTGAATC 657
 |||||
 Db 47 SerMetAlaGlyIleValGlnGlyThrAspLysLeuValPheIleAsnPhePheGlnIle 66
 QY 658 GATTTACATACGACCGCTGTCATGAGGAGAAAGACAGCAAGGCTTAAGTCAAAATCA 717
 |||||
 Db 67 AspLeuThrSerGlnHisAlaHisGlyLysThrGlnGlnLysLeuSerProLysSer 86
 QY 718 AAACATTTGCTACTGATAGTGGCGCATGTCATCAATTAACCTTGAAAGCTGACTTA 777
 |||||
 Db 87 LysProPheAlaThrAspAsnGlyAlaMetProHisLysLeuGlnLysAlaAspLeuLeu 106
 QY 778 AAGGCTATTCAGAACAAATGATCGTTACGCTCCACAGTACAGACTTATGAGGTC 837
 |||||
 Db 107 LysAlaIleGlnLysGlnLeuIleAlaAsnValHisSerAsnAspGlyIYrPheGlnVal 126
 QY 838 ATTGATTTTGCAGCGGATGCAACCATTAAGTATGCAAGCGGAGCTACTTTGCTGAC 897
 |||||
 Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyValIYrPheAlaAsp 146
 QY 898 AAAGATGCTTCGGTAACTTGCCGACCCCAACTGTCGAAGAATTTTGTAAAGCGGACAT 957
 |||||
 Db 147 LysAspGlySerValThrLeuProThrGlnProValGlnGlnPheLeuLeuSerGlyHis 166
 QY 958 GTGGCGTTAGACCATTAATAAGAAACCAATACAAACCAAGCAATCTGTTGATGTG 1017
 |||||
 Db 167 ValArgValArgProIYrLysGlnLysProValGlnAsnGlnAlaLysSerValAspVal 186
 QY 1018 GAATATATCTGATAGTTTACTCCCTTAACCTGATGACGATTTTCAAGACCGGCTCAA 1077
 |||||
 Db 187 LysIYrThrValGlnPheThrProLeuAsnProAspAspPheAspProGlyLysLeuLys 206
 QY 1078 GATATCAAGCTATTAAGAAACACTAGCATGCGGTGACACCATCATCATGCTCAAGAATTA 1137
 |||||
 Db 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeuLeu 226
 QY 1138 GCTCAAGCACAAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTATGAACGTGAC 1197
 |||||
 Db 227 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyIYrThrIleIYrGlnArgAsp 246
 QY 1198 TCGTCATGCTGCTACTGATGACATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 1257
 |||||
 Db 247 SerSerIleValIThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGln 266
 QY 1258 TTACTTACCGGTGTTAAATAACCGGAACAGCTTATGATCATTAATAAATCGGTCTG 1317
 |||||
 Db 267 PheThrIYrArgValLysAspArgGlnGlnAlaIYrGlyLysLeuLysSerGlyLeu 286
 QY 1318 AATGAAGAAATTAACAACTGACCTGATCTGTGAGAAATATTAAGTCTTAAATAAGG 1377
 |||||
 Db 287 AsnGlnGlnIleAsnAsnThrAspLeuIleSerGlnLysIYrThrIleLeuLysLysGly 306
 QY 1378 GAAAACCGTATGATCCCTTGTATGCGAGTCACTTGAACSTGTTCACCATCAATAGCTT 1437
 |||||
 Db 307 GluSerProIYrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysIYrVal 326
 QY 1438 GATGTCGATACCAACGAATTTGCTAAAAGTAGAGAGCTTTAAACAGCTACGAAAGCTAAC 1497
 |||||
 Db 327 AspValAsnThrAsnGlnLeuLeuLysSerGlnGlnLeuLeuThrAlaSerGlnArgAsn 346
 QY 1498 TTAGACTTCAGAGATTATACGATCCTCGATAGAGCTAACTACTCTCAACAATCTC 1557
 |||||
 Db 347 LeuAspPheArgAspLeuLysIYrAspProAlaGAspLysAlaLysLeuLeuLysIYrAsnAsnLeu 366
 QY 1558 GATGCTTTGTATTTATGACTATACCTTACTACGAAAGAGATAGAGATATACGATAC 1617
 |||||
 Db 367 AspAlaPheAspLleMetAspIYrThrLeuThrGlyLysValGlnAspAsnHisAspLys 386
 QY 1618 ACCAACCGTATCATTAACCGTTATATAGGCAACCGAACCGCAAGAGAGAAATGCTAGCTAT 1677

```

Db      387 AanaAsnArgIleValThrValIlyrMeGlyLysArgProLysGlyAlaLysGlySerTyr 406
Qy      1678 CATTAGCCTATGATTAAGAATCGTTATACCGAAGAAAGAAAGAGTTTACAGTACTG 1737
Db      407 HisLeuAlaIlyrAspLysAspLeuTyrThrGluGluGlnIlyrGlyAlaIlyrSerTyrLeu 426
Qy      1738 CGTTATACAGGACACCTTATACCTGATTAACCTTAACGACAAA 1779
Db      427 ArgAspThrGlnThrAlaIleProAspAsnProLysAspLys 440

RESULT 11
Q7X0Y7_STRPY
ID      Q7X0Y7_STRPY PRELIMINARY; PRT; 440 AA.
AC      Q7X0Y7;
DT      01-OCT-2003 (TrEMBLrel. 25, Created)
DT      01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT      01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE      Streptokinase.
GN      Name=ska;
OS      Streptococcus pyogenes.
OC      Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC      Streptococcus;
OC      NCB1_TaxID=1314;
RN      (1)
RP      NUCLEOTIDE SEQUENCE.
RC      STRAIN=IRP112;
RX      PubMed=14679231; DOI=10.1126/JB.186.1.110-121.2004;
RA      Kalia A., Beesen D.E.;
RT      "Natural selection and evolution of streptococcal virulence genes
RT      involved in tissue-specific adaptations.";
RL      J. Bacteriol. 186:110-121(2004).
DR      EMBL; AY234130; AAP3950.1; -, Genomic_DNA.
DR      HSSP; O53284; 1CAP.
DR      SMK; Q7X0Y7; 63-398.
DR      GO; GO:0016301; F:kinase activity; IEA.
DR      GO; GO:0008243; F:plasmaenogen activator activity; IEA.
DR      GO; GO:0005515; F:protein binding; IEA.
DR      InterPro; IPR004093; Staphylokinase.
DR      InterPro; IPR008124; Streptokinase.
DR      Pfam; PF02821; Staphylokinase; 3.
DR      PRINTS; PR01753; STREPKINASE.
KW      Kinase.
SQ      SEQUENCE 440 AA; 50160 MW; 1C5AF07907EC7AC8 CRC64;

Alignment Scores:
Pred. No.: 1,13e-125 Length: 440
Score: 1926.00 Matches: 368
Percent Similarity: 93.48% Conservative: 19
Best Local Similarity: 88.89% Mismatches: 27
Query Match: 61.46% Indels: 0
Db: 2 Gaps: 0

US-09-940-235-11 (1-1782) x Q7X0Y7_STRPY (1-440)
Qy      538 ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACGCAATTGGTTGT 597
Db      27 IleAlaGlyTyrGlyTyrPheLeuAspArgProProValAsnAsnSerGlnLeuVal 46
Qy      598 AGCGTTGCTGTAAGTCTTTGAGGGGAGCAATACAGACATTAGTCTTAAATTTTGAATC 657
Db      47 SerMetAlaGlyIleValGlnGlyThrAspLysLysValPheIleAsnPhePheGluIle 66
Qy      658 GATTAACATCAACGACCTGCTCATGAGAGAAAGACAGAGCCTTAAGTCCAAATCA 717
Db      67 AspLeuThrSerGlnProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 86
Qy      718 AAACCATTTGCTACTATAGTGGCGGAGTGTCAATTAATTTGAGAAAGTGAATTA 777
Db      87 LysProPheAlaThrAsnSerSerAlaMetProHisLysLeuGlnLysAlaAspLeu 106
Qy      778 AAGGCTATCAAGAAATTAATGATCGCTACGTCACAGTAACGACGACTTATGAGGTC 837

```

```

Db      107 LysAlaIleGlnGluArgLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal 126
Qy      838 ATTATATTGGCAACGATGCAACCATTAATGATCGAAACGGCAAGGTACTTTGGTCA 897
Db      127 IleAspPheAlaSerAspAlaThrIleThrAspHisGlnGlyLysValTyrPheAlaAsp 146
Qy      898 AAAGATGATTCGGTAACTTACCTTGCCGACCCGACCTGTCCAAAGATTTTGGTAAAGGACAT 957
Db      147 LysAspAspSerValThrLeuProThrGlnProValGlnGlnPheLeuAlaGlyHis 166
Qy      958 GTGCGCGTTAGACCATTAATAAGAAACCAATACAAACCAAGCAAAATCTGTGATGTG 1017
Db      167 ValArgValArgProTyrLysGluLysProIleGlnThrProAlaLysSerValAspVal 186
Qy      1018 GAATATCTGTGACGTTTACCTCCCTTAACCGCTGATGACGATTTTACGACCGAGTCAAA 1077
Db      187 ArgTyrThrValGlnPheThrProLeuAsnProAspAspPheAspProValLeuLys 206
Qy      1078 GATACATAAGCTATTTGAAAACACTAGCTACGGTGACACCATCATCTCAAGAAATTA 1137
Db      207 AsnThrLysLeuLeuLysThrLeuAlaIleGlyGlyThrValThrSerGlnGlnLeu 226
Qy      1138 GCTCAGCAACAGCATTTTAAACAAAACCCAGCGCTATACGATTATGAACGTGAC 1197
Db      227 AlaGlnAlaGlnSerIleLeuAsnGlnSerHisAspAspTyrThrIleTyrGluArgAsp 246
Qy      1198 TTCCATTCGTCATCTGACCAATGACATTTCCGTACGATTTTACCAATGATCAAGAG 1257
Db      247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGln 266
Qy      1258 TTTACTTACCGTGTAAAAATCGGAAACAGCTTATGATCAATTAATAAAATCGTGCTG 1317
Db      267 PheThrTyrThrIleLysIleLysAspArgGlnGlnAlaTyrGlyIleAsnLysSerGlyGln 286
Qy      1318 AATGAAGAAATTAACCAACCTGACCTGATCTTGAGAAATTAATGCTCTTAAAAAGG 1377
Db      287 GlnGlnLysThrAsnAsnThrAspLeuIleSerGlnLysTyrTyrValLeuLysGlyGly 306
Qy      1378 GAAAAGCGCTATGATCCCTTTGATCGCAGTCACTTGAACCTTGACCATTAATGATT 1437
Db      307 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleAsnTyrVal 326
Qy      1438 GATGCTGATCAACCAAGATTCCTAAAAAGTAGAGCGCTTAAACGCTAGGCAACGTAA 1497
Db      327 AspValAsnThrAsnLysLeuLysSerGlnGlnLeuLeuThrAlaSerGlnArgAsn 346
Qy      1498 TTAGACTTCAGAGATTATACGATCTCGTATAGGCTTAACTTACTTCAACATCTC 1557
Db      347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
Qy      1558 GATGCTTTGGTATTAATGACATTAACCTTAACTGGAAGAGTAAATACGATGAC 1617
Db      367 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspLys 386
Qy      1618 ACCAACCGTATCAATTAACGTTTATATGGGAGGACCCGAAAGAGGAATGCTAGTAT 1677
Db      387 AsnAsnArgIleValThrValIlyrMeGlyLysArgProGlnGlnLysAlaSerTyr 406
Qy      1678 CATTAGCCTATGATTAAGAATCGTTATACCGAAGAAAGAAAGAGTTTACAGTACTG 1737
Db      407 HisLeuAlaIlyrAspLysAspArgTyrThrGluGlnGlnIlyrGlyAlaIlyrSerTyrLeu 426
Qy      1738 CGTTATACAGGACACCTTATACCTGATTAACCTTAACGACAAA 1779
Db      427 ArgTyrThrGlyThrProIleProAspAsnProLysAspLys 440

RESULT 12
Q5X9T6_STRP6
ID      Q5X9T6_STRP6 PRELIMINARY; PRT; 440 AA.
AC      Q5X9T6;
DT      25-OCT-2004 (TrEMBLrel. 28, Created)
DT      25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT      25-OCT-2004 (TrEMBLrel. 28, Last annotation update)

```

DE Streptokinase (BC 3.4.-.-).
GN OrderedLocustNames=M6_Spy1692;
OS Streptococcus pyogenes (serotype M6).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=301450;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MGAS10394;
RX PubMed=15272401; DOI=10.1086/422697;
RA Banks D.J., Porcella S.F., Barbhan K.D., Beres S.B., Phillips L.R.,
Voyich J.M., Deleo F.R., Martin J.M., Somerville G.A., Musser J.M.;
RT "Progress toward characterization of the group A Streptococcus
RT metagenome: complete genome sequence of a macrolide-resistant serotype
RT M6 strain";
RL J. Infect. Dis. 190:727-738(2004).
DR EMBL; CP000003; AAT87827.1; -; Genomic_DNA.
DR SMR; OSX976; 63-398.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008243; F:plasmaenogen activator activity; IEA.
DR GO; GO:0005151; F:protein binding; IEA.
DR InterPro; IPR004093; Staphylokinase.
DR InterPro; IPR008124; Streptokinase.
DR Pfam; PF02821; Staphylokinase; 3.
DR PRINTS; PR01753; STREPKINASE.
KM Complete proteome; Hydrolase; Kinase.
SQ SEQUENCE 440 AA; 50160 MW; 1CSAF07907ECTAC8 CRC64;

Alignment Scores:
Pred. No.: 1.13e-125 Length: 440
Score: 1926.00 Matches: 368
Percent Similarity: 93.48% Conservative: 19
Best Local Similarity: 88.89% Mismatches: 27
Query Match: 61.46% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-11 (1-1782) x OSX976_STRP6 (1-440)

QY 538 ATTGTGACCTGAGTGGCTGTGCTAGACCGTTCATCTGTCAACAGCAAGCAATGTTGTT 597
DB 27 IleaIaGlyIyYrGlyIYrleuLeuLeuPaPrgProPovaIaAsnAsnSerGlnLeuVal 46
QY 598 AGCGTGTGCTGACTGTTGAGGGAGAGCATCAAGACATTAGCTTAAATTTTGAATC 657
DB 47 SerMetAlaGlyIleValGlnGlyIYrAspLysLysValPheIleAsnPheGluIle 66
QY 658 GATTCATCATCAAGCACTGCTCATGAGGAGAAAGACAGCAAGGCTTAAGTCAAAATGA 717
DB 67 AspLeuIlnrSerGlnProAlaHisGlyIYrLysIYrIlnGlnGlnGlyLeuSerProLysSer 86
QY 718 AAACCATTTGCTACTGATAGTAGCGCGCATGTCACTAACTTGAGAACTGACTTACTA 777
DB 87 LysPProPheAlaThrAsnSerSerAlaMetProHisLysLeuGlnLysIlaAspLeu 106
QY 778 AAGGTTATTCAGAAACATTTGATCGCTACGTCACAGTACAGCACTTGTGAGGTC 837
DB 107 LysAlaIleGlnGlnIYrAspLeuIleAlaAsnValHisSerAsnAspGlyIYrPheGluVal 126
QY 838 ATTGATTTTGAAGGAGATCAACCATTAATCGAAAGGGAAGGCTTACTTGTGTCAG 897
DB 127 IleAspPheAlaSerAspAlaThrIleIlnrAspAsnGlyLysValIYrPheAlaAsp 146
QY 898 AAAGATGGTTCGTAACCTTGGCGAACCAACCTGTCTCAAGATTTTGTGAAGCGACAT 957
DB 147 LysAspAspSerValIThrLeuProIlnrGlnProValGlnGlnPheLeuLeuArgIYrHis 166
QY 958 GTGGCGCTTGAACCATATTAAGAAAAACAATACAAACCAAGCAAGCAATCTGTTGATG 1017
DB 167 ValArgValArgProIYrLysGlnLysProIleGlnIlnrProAlaLysSerValAspVal 186
QY 1018 GAATTAATCTGTAAGATTTACTCCCTTAACCTCGATGAGATTTTGACAGAGCTTCANA 1077

DB 187 ArgTyIThrValGlnPheThrProLeuAsnProAspAspPheArgProValIleuLys 206
QY 1078 GATACCTAAGCTATTTGAAAAACATAGCTATGGTGACACCATCATCTTCAAGATTACTA 1137
DB 207 AsnThrLysLeuLeuLysThrIleuAlaIleGlyGlyIYrValIlnrSerGlnGlnLeu 226
QY 1138 GCTCAAGCAACAAAGCATTTTAAACAAAACACACCGAGCTATAGATTTTGAACGTGAC 1197
DB 227 AlaGlnIlnrAlaSerIleuLeuAsnGlnSerHisProAspIYrThrIleIYrGlnValAsp 246
QY 1198 TCCTCAATCGTCACTGATGACAAATGACATTTTCCGTAAGATTTTCAACATGATCAAG 1257
DB 247 SerSerIleValIlnrHisAspAsnAspIlePheArgThrIleuProMetAlaAspGlnGln 266
QY 1258 TTTACTTACCGGTTTTAAAAATCGGGAACAGCTTATAGATCAATAAAAATCTGCTG 1317
DB 267 PheThrTyIlnrHisLeuLysAspArgGlnAlaTyIleAsnLysLysSerGlyGln 286
QY 1318 AATGAAGAAATAAACAACATGACCTGATCTGAGAAATTTACGCTTAAAAAAAGG 1377
DB 287 GlnGlnLysThrAsnIlnrAspLeuIleSerGlnLysTyIYrValIleuLysLysGly 306
QY 1378 GAAAGCCGATGATGATCCCTTTGATCGCAGTCACTTGAACCTGTCACATCAATACGTT 1437
DB 307 GlnLysProIYrAspProPheAspArgSerHisLeuLysLeuPheThrIleAsnTyIVal 326
QY 1438 GATGTGATACCAACGATTTGCTAAAGAGTGAAGAGCTTTAACAGCTTACGGAACGTAAC 1497
DB 327 AspValAlaAsnThrAsnLysLeuLysSerGlnGlnLeuLeuThrAlaSerGlnValArgAsn 346
QY 1498 TTGACCTGACAGATTTTATACGATCCCGGATTAAGGCTTAAAGCTTACATCTCAACAATCTC 1557
DB 347 LeuAspPheArgAspLeuTyIAspProArgAspLysAlaLysLeuLeuTyIAsnAsnLeu 366
QY 1558 GATGCTTTTGGTATTTATGACATATACCTTAACTGGAAGAAAGTAGAGTAATCAAGATAC 1617
DB 367 AspAlaPheGlyIleMetAspTyIThrIleuThrGlyLysValGlnAspAsnHisAspLys 386
QY 1618 ACCAACCGTATCATTAACCGTTATATAGGAGACGACCCGAGAGAGAAATGCTAGTAT 1677
DB 387 AsnAsnArgValValThrValTyIYrMetGlyLysArgProGlnGlnAlaAsnLysTyI 406
QY 1678 GATTTAGCTTATGATTAAGATCGTTATACCGAAGAAAGAAAGATTACAGTACTAC 1737
DB 407 HisLeuAlaIYrAspLysAspArgTyIYrGlnGlnGlnLysValIYrSerTyILeu 426
QY 1738 CGTTATACAGGACCACTATACCTGATTAACCTTAACGACAA 1779
DB 427 ArgTyIThrGlyIYrThrProIleProAspAsnProLysAspLys 440

RESULT 13
Q8NZAG_STRP8
ID Q8NZAG_STRP8 PRELIMINARY; PRT; 440 AA.
AC Q8NZAG6;
IC Q8NZAG6;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Streptokinase.
GN OrderedLocustNames=spyM18_2042;
OS Streptococcus pyogenes (serotype M18).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=301451;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MGAS8232 / Serotype M18;
RX MEDLINE=21927593; PubMed=11917108; DOI=10.1073/pnas.062526099;
RA Smoot J.C., Barbhan K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
Sylvia G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,
RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
RT "Genome sequence and comparative microarray analysis of serotype M18
RT group A Streptococcus strains associated with acute rheumatic fever

RT outbreake.",
 RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673 (2002).
 DR EMBL; AB010107; AL98517.1; -; Genomic_DNA.
 DR HSSP; O53284; 1C4P.
 DR SMR; O8NZ6; 63-398.
 DR GO; GO:0016301; F:kinase activity; IEA.
 DR GO; GO:0008243; F:plasmaenogen activator activity; IEA.
 DR GO; GO:0005515; F:protein binding; IEA.
 DR InterPro; IPR004093; Staphylokinase.
 DR Pfam; PF02821; Staphylokinase; 3.
 DR PRINTS; PR01753; STREPKINASE.
 KW Complete proteome; Kinase.
 SQ SEQUENCE 440 AA; 50382 MW; 47DAB35577EBB88B CRC64;

Alignment Scores:

Pred. No.:	1.83e-125	Length:	440
Score:	1923.00	Matches:	364
Percent Similarity:	93.96%	Conservative:	25
Best Local Similarity:	87.92%	Mismatches:	25
Query Match:	61.36%	Indels:	0
DB:	2	Gaps:	0

US-09-940-235-11 (1-1782) x Q8NZ6_STRP8 (1-440)

QY 538 ATGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCAATGGTGT 597
 DB 27 TLeaIGlYrGluTtrPLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 46
 QY 598 AGCGTGTCTGTACTGTGAGGAGCAATCAAGACATTAAGTCTTAATTTTGAATC 657
 DB 47 SerMetAlGlyLeuValGluGluGlyThrAspLeuValPheIleAsnPheGlnIle 66
 QY 658 GATCTAACATCAGACCTGTCTATGAGGAAAGACAGCAAGGCTTAAGTCCAAATCA 717
 DB 67 AspLeuThrSerInProAlaHisGlyGlyLeuThrGluGluGlyLeuSerProLysSer 86
 QY 718 AAACGATTTGCTACTGATAGTGGCGGATGTCAATTAACCTTGAGAAAGCTTACTA 777
 DB 87 LysProPheAlaThrAspAsnGlyAlaMetProHisLysLeuGluValAspLeuLeu 106
 QY 778 AAGCTATTCAGAACATTTGATCGCTAACGTCACAGTACGACTATTGAGATC 837
 DB 107 LysAlaIleGlnIleArgLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal 126
 QY 838 ATTGATTTTGCAGACGATGCAACCATTAAGTATGATGAAAGCGCAAGTCTTACTGAC 897
 DB 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAspAsnIleTyrPheAlaAsn 146
 QY 898 AAAGATGTTGCGGTACCTTGCCGACCCGACCTGTCCAAAGATTTTGTGAAGGAT 957
 DB 147 GlnAspGlySerValThrLeuProLysGlnProIleGlnIlePheLeuLeuAlaGlyHis 166
 QY 958 GTGCGGCTTGAACCATTAAGAAACCAATACAAACCAAGGGAATCGTTGAGATG 1017
 DB 167 ValArgValArgProTyrLysGluLysProIleGlnIleThrProAlaLysSerValAspIle 186
 QY 1018 GAATATACGTATACGTTTACCTCCCTTAAACCTGATGACGATTTTCAGACCAAGTCTCAA 1077
 DB 187 ArgTyrThrValGlnPheThrProLeuAsnProAspAspAspPheLysProValLeuLys 206
 QY 1078 GATACCTAGCTATTTGAAACCACTAGCTATGGTGACCAATCAGATCTCAAGATTA 1137
 DB 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnIleLeuLeu 226
 QY 1138 GGTCAAGCAACAAAGCTTTTAAACAAACCAACCAAGGCTATCGATTTTGAACGTCAC 1197
 DB 227 AlaGlnAlaGlnSerIleLeuAsnGlnSerHisSerAspTyrThrIleTyrGluAspAsp 246
 QY 1198 TCTCATGCTGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1257
 DB 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnIle 266

QY 1258 TTTACTTACCGGTGTTAAAAATCGGAGACAGCTTATAGATCATATAAAAAATCTGCTG 1317
 DB 267 PheThrTyrHisIleLysAspArgGluGlnAlaTyrGlyIleAsnLysLysSerIleGln 286
 QY 1318 AATGAAAGAAATTAACACACATGACCTGATCTCTGAGAAATATTAAGTCTTAAAAAGG 1377
 DB 287 GluGluLysThrAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysGly 306
 QY 1378 GAAAGCCGTATGATCCCTTTGATCGGAGTCACTTGAACCTTTACCATCAAAATCGTT 1437
 DB 307 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleAsnTyrVal 326
 QY 1438 GATGCTGATACCAAGAAATGCTAAAGAGTGAAGCTCTTACAGCTAGCAAGCTAAC 1497
 DB 327 AspValAsnThrAsnLysLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
 QY 1498 TTGACCTTCAGAGATTTATACGATCTGTGATTAAGGTTAACTATCTACCAATCTC 1557
 DB 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
 QY 1558 GATGCTTTGTTATATGACCTATACCTTAAGTGAAGAAAGTATGAGATTAATCAGATGAC 1617
 DB 367 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspLys 386
 QY 1618 ACCAACCGTATCATACCGTTTATATGAGGACGACCGGAGAGAGAAATGCTAGCTAT 1677
 DB 387 AsnAsnArgValAlaThrValTyrMetGlyLysArgProGluGluGluAsnAlaSerTyr 406
 QY 1678 CATTAGCGCTATGATTAAGATCGTTATACCGAAGAAAGCAAGATTAAGCTACCTG 1737
 DB 407 HisLeuAlaTyrAspLysAspArgTyrThrGluGluGluValTyrSerTyrLeu 426
 QY 1738 CGTTATACAGGACACTATACCTGATACCTTAACCTTAACGACAA 1779
 DB 427 ArgTyrThrGlyThrProIleProAspAsnProLysAspLys 440

RESULT 14

Q7X0Y3_STRPY
 ID Q7X0Y3_STRPY PRELIMINARY; PRT; 440 AA.
 AC Q7X0Y3;
 DT 01-OCT-2003 (TREMBlrel. 25, Created)
 DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE Streptokinase.
 GN Name=aka;
 OS Streptococcus pyogenes.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1314;
 RX [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=ALAB49;
 RX PubMed=14679231; DOI=10.1128/JB.186.1.110-121.2004;
 RA Kalia A., Beasen D.B.;
 RT "Natural selection and evolution of streptococcal virulence genes
 involved in tissue-specific adaptations.";
 RT J. Bacteriol. 186:110-121(2004).
 RL EMBL; AY234134; AAP39954.1; -; Genomic_DNA.
 DR HSSP; O53284; 1C4P.
 DR SMR; O7X0Y3; 63-398.
 DR GO; GO:0016301; F:kinase activity; IEA.
 DR GO; GO:0008243; F:plasmaenogen activator activity; IEA.
 DR GO; GO:0005515; F:protein binding; IEA.
 DR InterPro; IPR004093; Staphylokinase.
 DR Pfam; PF02821; Staphylokinase; 3.
 DR PRINTS; PR01753; STREPKINASE.
 KW Kinase.
 SQ SEQUENCE 440 AA; 50342 MW; B3C6DBAA50DB2C39 CRC64;

Alignment Scores:

Pred. No.:	1.82e-122	Length:	440
Score:	1880.00 <td>Matches: <td>356</td> </td>	Matches: <td>356</td>	356

Percent Similarity: 92.51%
 Best Local Similarity: 85.99%
 Query Match: 59.99%
 DB: 2
 Conservative: 27
 Mismatches: 31
 Indels: 0
 Gaps: 0

US-09-940-235-11 (1-1782) x Q7X0Y3_STRPY (1-440)

```

QY 538 ATTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCAATTGGTTGT 597
DB 27 ILeAlaGlyTyrGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 46
QY 598 AGCGTGTCTGTACTGTGAGGGAGCAATCAAGACATTAGTCTAAATTTTGAATC 657
DB 47 SerMetAlaGlyIleValGlnGlyThrAspIleValPheIleAsnPhePheGluIle 66
QY 658 GATCTTAACATCAAGACCTGTCTAGAGAAAGACAGAGAGCTTAAGTCCAAATCA 717
DB 67 AspLeuThrSerGlnProAlaHisGlyGlyLeuThrGlnGlnGlyLeuSerProIleSer 86
QY 718 AAACCATTTGCTACTGATAGTGGCGAGTGTCACTTAACCTTGAGAAGCTGACTTA 777
DB 87 LysProPheAlaThrAspAsnGlyAlaMetProHisLeuGlnGlyValAspLeu 106
QY 778 AAGGCTATTCAAGAACATTTGATGCTTAACGTCACAGTAAACGACGACTTTTGAAGTC 837
DB 107 LysAlaIleGlnGlnLeuLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGlnVal 126
QY 838 ATTGATTTTGCAGAGCATGACCACTTACTGATGCAAGACGCGACTTACTTGTCTGAC 897
DB 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAspAsnIleTyrPheAlaAsn 146
QY 898 AAAGATGTTGGGTACTGCTTGGCCGACCCGACCTGTCCAGAAATTTTGTCAACGCAAT 957
DB 147 GlnAspGlySerValThrLeuProGlnGlnProIleGlnGlnLeuLeuValGlyHis 166
QY 958 GTGGCGGTGAGACCATATTAAGAAAAACCAATCAAAACCAAGCAAGTCTGTGATGTC 1017
DB 167 ValAlaGlyValArgProIleTyrGlnGlySerProIleGlnThrProAlaIleSerValAspIle 186
QY 1018 GAATTAATCTGTACAGTTACTCCCTTAAACCTGTAGTACGATTTTCAGACGAGCTTCA 1077
DB 187 ArgTyrThrValGlnPheThrProLeuAsnProAspAspPheLysProValLeuLys 206
QY 1078 GATTAATCACTTATGAAACACTAGTATCGGTGACACATCACTCAAGTCAAGAAATTA 1137
DB 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeuLeu 226
QY 1138 GCTCAAGACCAAAAGCATTTTAAACAACAAACCCAGGCTATGATTTTGAACGCTGAC 1197
DB 227 AlaGlnIleGlnSerIleLeuAsnGlnSerHisProAspTyrThrIleTyrGlnIleArgAsp 246
QY 1198 TCCTCAATGCTCATCATGACCAATGACATTTTCGTAAGATTTTACCAATGATCAAG 1257
DB 247 SerSerIleValIleThrHisAspAsnAspIlePheArgThrIleLeuProThrArgGlnGlu 266
QY 1258 TTTACTTACCGTGTAAATAATGGGAAACAAGCTTATAGATCAATAATAATCGCTGTC 1317
DB 267 PheThrTyrHisIleValIleAsnArgGlnGlnIleAlaTyrGlnAsnAspAsnIleTyrGlnLeu 286
QY 1318 AATGAAGAATAAACAACACTGACCTGATCTGTGAATAATATATATATATATATATAT 1377
DB 287 LysLysGlnThrLysAsnThrAspLeuIleSerGlnIleTyrTyrIleLeuLysLysGly 306
QY 1378 GAAAAGCCGTATGATCCCTTGTATGCGAGTCACTTGAATCTGTCAACATCAATAGCTT 1437
DB 307 GlnLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326
QY 1438 GATGTGATCAACGAATTTGCTTAAATAAGTAGACAGCTTTTAAAGCTAGCAAGCTAAC 1497
DB 327 AspValAspThrLysAspLeuLysSerGlnGlnIleLeuThrIleAspGlnIleAsn 346
QY 1498 TTAGACTTCAGAGATTATACGATCTCGTATTAAGCTTAACTACTCTTCAACAATATTC 1557

```

```

DB 347 LeuAspPheArgAspLeuTyrAspProArgAspIleAlaLysLeuLeuTyrAsnAsnLeu 366
QY 1558 GATGCTTTGGTATTATAGCACTTAACCTTAACTGGAAGAAAGTAAATCAAGATGAC 1617
DB 367 AspAlaPheAspIleMetAspTyrThrLeuThrGlyLysValGlnAspAsnHisAspLys 386
QY 1618 ACCAACCGTATCAATTAACCGTTTATATAGGCAAGCAACCCGAAGAGAGAATGCTAGCTAT 1677
DB 387 AsnAsnArgValValThrValTyrMetGlyLysArgProLysGlyAlaLysGlySerTyr 406
QY 1678 CATTTAGCTTATGATTAAGATCGTTATACCGAAGAACGAGAAAGTTTACGCTACTG 1737
DB 407 HisLeuAlaTyrAspLysAspLeuTyrThrGlnGlnGlnIleValTyrSerTyrLeu 426
QY 1738 CGTTATACAGGACGACCTTATACCTGATTAACCTTACGACGAA 1779
DB 427 ArgAspThrGlyThrProIleProAspAsnProAsnAspLys 440

RESULT 15
Q7X0Y0_STRPY PRELIMINARY; PRT; 440 AA.
AC Q7X0Y0;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Streptokinase.
GN Name=ska;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=d633;
RX PubMed=14679231; DOI=10.1128/JB.186.1.110-121.2004;
RA Kalia A., Bessen D.B.;
RT "Natural selection and evolution of streptococcal virulence genes
RT involved in tissue-specific adaptations.";
RL J. Bacteriol. 186:110-121(2004).
DB EMBL; AY234138; AAP39958.1; -, Genomic_DNA.
DR HSSP; Q53284; 1CAP.
DR SMK; Q7X0Y0; 63-398.
DR GO; GO:0016301; F-kinase activity; IEA.
DR GO; GO:0008243; F-platelet activator activity; IEA.
DR GO; GO:0005515; F-protein binding; IEA.
DR InterPro; IPR004093; Streptokinase.
DR Pfam; PF02821; Streptokinase; 3.
DR PRINTS; PRO1753; STREPKINASE.
KW kinase.
SQ SEQUENCE 440 AA; 50143 MW; 4ACB2C29349D680C CRC64;

Alignment Scores:
Pred. No.: 6,58e-122 Length: 440
Score: 1872.00 Matches: 354
Percent Similarity: 92.03% Conservative: 27
Best Local Similarity: 85.51% Mismatches: 33
Query Match: 59.73% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-11 (1-1782) x Q7X0Y0_STRPY (1-440)
QY 538 ATTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCAATTGGTTGT 597
DB 27 ILeAlaGlyProGluTrpLeuLeuGlyArgProValAsnAsnSerGlnLeuVal 46
QY 598 AGCGTGTCTGTACTGTGAGGGAGCAATCAAGACATTAGTCTAAATTTTGAATC 657
DB 47 SerMetAlaGlyIleValGlnGlyThrAspIleValPheIleAsnPhePheGluIle 66
QY 658 GATCTTAACATCAAGACCTGTCTAGAGAAAGACAGAGAGCTTAAGTCCAAATCA 717
DB 67 AspLeuThrSerGlnProAlaHisGlyGlyLeuThrGlnGlnGlyLeuSerProIleSer 86

```



```

Db      1369 AAAGGGAAGTCTAAGGCTATTCAGAACTATGATCCCTAACGCTCCACAGTAACGAC 1428
Qy      873 GACTACTTTGAGGCTCATGATTTTGGCAAGGATGCAACCTTACTGATCCGAAACGGCAAG 932
Db      1429 GACTACTTTGAGGCTCATGATTTTGGCAAGGATGCAACCTTACTGATCCGAAACGGCAAG 1488
Qy      933 GCTTACTTTGCTGCAAGAAAGTGGTGGTAACTGGCCGACCCCAACCTGTCGAAGATTT 992
Db      1489 GCTTACTTTGCTGCAAGAAAGTGGTGGTAACTGGCCGACCCCAACCTGTCGAAGATTT 1548
Qy      993 TTGCTAAGCGGACATGTCGCGCTTGAACCATATTAAGAAAAACCATATCAAAACCAACG 1052
Db      1549 TTGCTAAGCGGACATGTCGCGCTTGAACCATATTAAGAAAAACCATATCAAAACCAACG 1608
Qy      1053 AAATCTGTTGATGATGCAATATGATGATGATGATGATGATGATGATGATGATGATGAT 1112
Db      1609 AAATCTGTTGATGATGCAATATGATGATGATGATGATGATGATGATGATGATGATGAT 1668
Qy      1113 AAGCGAGGCTGCAAGAAAGTGGTGGTAACTGGCCGACCCCAACCTGTCGAAGATTT 1172
Db      1669 AAGCGAGGCTGCAAGAAAGTGGTGGTAACTGGCCGACCCCAACCTGTCGAAGATTT 1728
Qy      1173 TCTCAAGATTTACTAGCTCAGCAACAAGCATTTTAAACAAAACCAACGAGCTATACG 1232
Db      1729 TCTCAAGATTTACTAGCTCAGCAACAAGCATTTTAAACAAAACCAACGAGCTATACG 1788
Qy      1233 ATTTATGAAGCTGATCTCTCAATGCTCATGATGATGATGATGATGATGATGATGATGAT 1292
Db      1789 ATTTATGAAGCTGATCTCTCAATGCTCATGATGATGATGATGATGATGATGATGATGAT 1848
Qy      1293 CCAATGCAATCAAGATTTACTAGCTCAGCAACAAGCATTTTAAACAAAACCAACGAGCTAT 1352
Db      1849 CCAATGCAATCAAGATTTACTAGCTCAGCAACAAGCATTTTAAACAAAACCAACGAGCTAT 1908
Qy      1353 AAAAAAATCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1412
Db      1909 AAAAAAATCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1968
Qy      1413 GTCTTTAAAAAAGGGGAAAAAGCGTATGATCCCTTTGATGCAAGTCACTTGAACCTGTC 1472
Db      1969 GTCTTTAAAAAAGGGGAAAAAGCGTATGATCCCTTTGATGCAAGTCACTTGAACCTGTC 2028
Qy      1473 ACCATCAAAATAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1532
Db      2029 ACCATCAAAATAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2088
Qy      1533 GCTAGGGAAGTAACTTAACTTCAAGATTTTAAAGATTTTAAAGATTTTAAAGATTTTAA 1592
Db      2089 GCTAGGGAAGTAACTTAACTTCAAGATTTTAAAGATTTTAAAGATTTTAAAGATTTTAA 2148
Qy      1593 CTCTACAAACATCTCGATGCTTTTGGTATTTAGGATCTTAACTTGAAGAAAGTATGAG 1652
Db      2149 CTCTACAAACATCTCGATGCTTTTGGTATTTAGGATCTTAACTTGAAGAAAGTATGAG 2208
Qy      1653 GATTAATCAAGATGACCAACCGTATATACCGTTTATATGGGCAAGCAACCCGAAGGA 1712
Db      2209 GATTAATCAAGATGACCAACCGTATATACCGTTTATATGGGCAAGCAACCCGAAGGA 2268
Qy      1713 GAGAAATGCTAGCTACATTTAGCTGGTGT 1742
Db      2269 GAGAAATGCTAGCTACATTTAGCTGGTGT 2298

```

```

RESULT 2
STRSC   STRSC      2568 bp   DNA      linear   BCT 26-Apr-1993
DEFINITION Streptococcus equisimilis (H46A) streptokinase gene, complete cds.
ACCESSION K02986
VERSION   K02986.1 GI:153808
KEYWORDS  streptokinase.
SOURCE    Streptococcus dysgalactiae subsp. equisimilis
ORGANISM  Streptococcus dysgalactiae subsp. equisimilis

```

```

Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.
REFERENCE
1 (bases 1 to 2568)
AUTHORS
Malke,H., Roe,B. and Ferretti,J.J.
TITLE
Nucleotide sequence of the streptokinase gene from Streptococcus
equisimilis H46A
JOURNAL
PUBMED
Gene 34 (2-3), 357-362 (1985)
2989113
COMMENT
Original source text: S. equisimilis (strain H46A) DNA, clone pMF5.
Draft entry and hard copy of sequence for [1] kindly provided by
J.J.Ferretti, 03-SEP-1985.
The -35 and -10 regions are located at positions 760-765 and
781-786 respectively and an SD sequence at 808-813. Downstream
from the coding region inverted repeats (positions 2176-2190 and
2203-2217) are thought to function as transcription terminators.
The nucleotide sequence of sic does not support the hypothesis that
the gene has evolved by duplication and fusion, as suggested by
internal two-fold AA homologies of its product.
FEATURES
source
location/Qualifiers
1..2568
/organism="Streptococcus dysgalactiae subsp. equisimilis"
/mol_type="genomic DNA"
/db_xref="taxon:119602"
794..>2141
/product="skc mRNA"
819..2141
/note="pre-streptokinase"
/codon_start=1
/translation="MKNYLSFGMALLPALTFGVNVOATAGBEMLLDRSVNNSOL
VSVAGTVEGNDISLKFPEIDITSPPHAGKTEQGLSPSKFPATDSGAMSHLEK
ADILKAIQEOILANVHNDYFEVIDASATQITDRGKVFAPKDSVTLPTQVOR
FLLSGHVRVPYRKPIONOKASVDEVTQFTPLNDDPFRPLKDTKLKTLAIGD
TTSOELLAOOSILANKHGYTLYEEDSSIVTHDNDIPRILPMDOEFYRVNREO
AVRINKSGINERINNTDILSEKYYVYKGGKPYDPDRSLKIFTKYVVDVNNELL
KSEDLTLASERNLDRDLYPDRRAKLLYNLDFGMDYTLTGKVEDNHDNTRIIIT
VYMKRPEGENASVHLYVDKRYTEEREVSYLYRYGTPIPNPNDK"
819..896
/note="streptokinase signal peptide"
mat_peptide
897..2138
/product="streptokinase"
ORIGIN
5 bp upstream of Pect site.
Query Match 54.7%; Score 1146.4; DB 1; Length 2568;
Best Local Similarity 99.5%; Pred. No. 1.2e-297;
Matches 1150; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

```

Qy      587 TATTGCTGACCTGAGTGGCTGTAGACCGTCACTGTGCAACAACGCCAATTGTTGT 646
Db      896 TATTGCTGACCTGAGTGGCTGTAGACCGTCACTGTGCAACAACGCCAATTAGTTGT 955
Qy      647 TAGGTTGCTGATCTGTGAGGGGAGCAATCAAGACATTAAGTCTTAAATTTTGAAT 706
Db      956 TAGGTTGCTGATCTGTGAGGGGAGCAATCAAGACATTAAGTCTTAAATTTTGAAT 1015
Qy      707 CGATCTAACATCAGCACTGCTCATGAGGAAAGACAGAGCAAGCTTAAAGTCCAAATC 766
Db      1016 CGATCTAACATCAGCACTGCTCATGAGGAAAGACAGAGCAAGCTTAAAGTCCAAATC 1075
Qy      767 AAAACCAATTTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 826
Db      1076 AAAACCAATTTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1135
Qy      827 AAAGGCAATTCAGAAACATTTGATCGCTAAGCTCAACAGTAAACGATCTTGGAGT 886
Db      1136 AAAGGCAATTCAGAAACATTTGATCGCTAAGCTCAACAGTAAACGATCTTGGAGT 1195
Qy      887 CATGATTTTGAAGCAAGCAATTAAGTCAAGAAACGCAAGGCTCTACTTTGCTGA 946
Db      1196 CATGATTTTGAAGCAAGCAATTAAGTCAAGAAACGCAAGGCTCTACTTTGCTGA 1255

```

QY 947 CAAAGATGTTCCGTAACTTCCGACCCCAACTCTGTCCAAAGATTTTCTTAAGCGACA 1006
| | | | |
Db 1256 CAAAGATGTTCCGTAACTTCCGACCCCAACTCTGTCCAAAGATTTTCTTAAGCGACA 1315
| | | | |
QY 1007 TGTGCGGTTAGACCAATATTAAGAAAAACAATTAACAACCAAGCGAATCTGTGTAGT 1066
| | | | |
Db 1316 TGTGCGGTTAGACCAATATTAAGAAAAACAATTAACAACCAAGCGAATCTGTGTAGT 1375
| | | | |
QY 1067 GGAATATCTGTACAGTTTACTCCCTTAAACCCGTATGACGATTTGACACGAGTCTCAA 1126
| | | | |
Db 1376 GGAATATCTGTACAGTTTACTCCCTTAAACCCGTATGACGATTTGACACGAGTCTCAA 1435
| | | | |
QY 1127 AGATATTAAGCTATTGAAAACTAGTATCGGTGACACATCAATCTCAAGAAATTA 1186
| | | | |
Db 1436 AGATATTAAGCTATTGAAAACTAGTATCGGTGACACATCAATCTCAAGAAATTA 1495
| | | | |
QY 1187 AGCTCAAGACAAAGATTTTAAACAACCAACCCGCTATACGATTTATGAAAGTGA 1246
| | | | |
Db 1496 AGCTCAAGACAAAGATTTTAAACAACCAACCCGCTATACGATTTATGAAAGTGA 1555
| | | | |
QY 1247 CTCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGATCAAGA 1306
| | | | |
Db 1556 CTCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGATCAAGA 1615
| | | | |
QY 1307 GTTTACTTACCGTGTAAAAATCGGAAACAAGCTTATAGATCAATAAAAATCTGTCT 1366
| | | | |
Db 1616 GTTTACTTACCGTGTAAAAATCGGAAACAAGCTTATAGATCAATAAAAATCTGTCT 1675
| | | | |
QY 1367 GAATGAAGAAATTAACAACACTGACCTGATCTGAGAAATATTAAGTCTTTAAAAAGG 1426
| | | | |
Db 1676 GAATGAAGAAATTAACAACACTGACCTGATCTGAGAAATATTAAGTCTTTAAAAAGG 1735
| | | | |
QY 1427 GGAAGAGCGTATGATCCCTTGTATGAGAGTCACTTGAACCTGTGCAACATTAAGT 1486
| | | | |
Db 1736 GGAAGAGCGTATGATCCCTTGTATGAGAGTCACTTGAACCTGTGCAACATTAAGT 1795
| | | | |
QY 1487 TGAATGATACCAAGAAATGCTAAAAAGTGAAGCTCTTAAACAGTACGACGAACTTAA 1546
| | | | |
Db 1796 TGAATGATACCAAGAAATGCTAAAAAGTGAAGCTCTTAAACAGTACGACGAACTTAA 1855
| | | | |
QY 1547 CTTAGACTTCAAGATTTTATAGATCTCTGTGATTAAGGCTTAACTACTTCAACAAATCT 1606
| | | | |
Db 1856 CTTAGACTTCAAGATTTTATAGATCTCTGTGATTAAGGCTTAACTACTTCAACAAATCT 1915
| | | | |
QY 1607 GGAATGTTTGGATTAATGACATTAACCTTAACTGAAAAATGAGATATCAAGATGA 1666
| | | | |
Db 1916 GGAATGTTTGGATTAATGACATTAACCTTAACTGAAAAATGAGATATCAAGATGA 1975
| | | | |
QY 1667 CACCAACCGTATCATACCGTTTATATGCGCAAGCGACCCGAAGAGAGATGCTAGCTA 1726
| | | | |
Db 1976 CACCAACCGTATCATACCGTTTATATGCGCAAGCGACCCGAAGAGAGATGCTAGCTA 2035
| | | | |
QY 1727 CCATTTAGCTGTGCT 1742
| | | | |
Db 2036 TCATTTAGCTATGAT 2051
| | | | |

RESULT 3
A04926 2568 bp DNA linear PAT 15-JUL-1993
LOCUS S.equisimilis skc gene for streptokinase.
DEFINITION A04926
ACCESSION A04926.1 GI:412219
VERSION A04926.1
KEYWORDS streptokinase.
SOURCE Streptococcus dysgalactiae subsp. equisimilis
ORGANISM Streptococcus dysgalactiae subsp. equisimilis
Bacteria; Firmicutes; Lactobacilliales; Streptococcaceae;
Streptococcus.
1 (bases 1 to 2568)
REFERENCE
AUTHORS Ferretti, J.J. and Malke, H.
TITLE Streptokinase-coding recombinant vectors
JOURNAL Patent: EP 0151337-A 1 14-AUG-1985;
THE BOARD OF REGENTS OF THE UNIVERSITY OF OKLAHOMA

FEATURES
source
Location/Qualifiers
1..2568
/organism="Streptococcus dysgalactiae subsp. equisimilis"
/mol_type="unassigned DNA"
/sub_species="equisimilis"
/db_xref="taxon:119602"
ORIGIN
Query Match 54.7%; Score 1146.4; DB 6; Length 2568;
Best Local Similarity 99.5%; Pred. No. 1.2e-297;
Matches 1150; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 587 TATTGCTGACCTGAGTGGCTGCTAGACCGTCACTGTCTCAACAACGCAATGTTGT 646
| | | | |
Db 896 TATTGCTGACCTGAGTGGCTGCTAGACCGTCACTGTCTCAACAACGCAATGTTGT 955
| | | | |
QY 647 TAGGTTGCTGTGATCTGTGAGGGGACAAATCAAGACATTAGCTTTAAATTTTGAAT 706
| | | | |
Db 956 TAGGTTGCTGTGATCTGTGAGGGGACAAATCAAGACATTAGCTTTAAATTTTGAAT 1015
| | | | |
QY 707 CGATCTAACATCAGACCTGCTCATGAGGAAAGACAGAGCAAGCTTAAGTCCAAATC 766
| | | | |
Db 1016 CGATCTAACATCAGACCTGCTCATGAGGAAAGACAGAGCAAGCTTAAGTCCAAATC 1075
| | | | |
QY 767 AAAACCAATTTGCTACTGATAGTGGCGCATGTCAATTAACCTTGAGAAAGCTGACT 826
| | | | |
Db 1076 AAAACCAATTTGCTACTGATAGTGGCGCATGTCAATTAACCTTGAGAAAGCTGACT 1135
| | | | |
QY 827 AAAGGCTATTCAGAAACAATGATGCTTAACGTTCAAGTACAGTAAAGCTTAAAGGT 886
| | | | |
Db 1136 AAAGGCTATTCAGAAACAATGATGCTTAACGTTCAAGTACAGTAAAGCTTAAAGGT 1195
| | | | |
QY 887 CATGATTTTGGCAAGCCATGCAACCATTACTGATTCGAAACGCGCAAGTCTACTTGTCTGA 946
| | | | |
Db 1196 CATGATTTTGGCAAGCCATGCAACCATTACTGATTCGAAACGCGCAAGTCTACTTGTCTGA 1255
| | | | |
QY 947 CAAAGATGTTGCGTAACTTGGCCGACCCCACTGTCCAAAGATTTTGTCTAAGCGGACA 1006
| | | | |
Db 1256 CAAAGATGTTGCGTAACTTGGCCGACCCCACTGTCCAAAGATTTTGTCTAAGCGGACA 1315
| | | | |
QY 1007 TGTGCGGCTTAAACATATTAAGAAAAACAATTAACAACCAAGCGAATCTGTGTAGT 1066
| | | | |
Db 1316 TGTGCGGCTTAAACATATTAAGAAAAACAATTAACAACCAAGCGAATCTGTGTAGT 1375
| | | | |
QY 1067 GGAATTAATCTGATCAATTTACTCCCTTAAACCTGTATGACGATTTGACACGAGTCTCAA 1126
| | | | |
Db 1376 GGAATTAATCTGATCAATTTACTCCCTTAAACCTGTATGACGATTTGACACGAGTCTCAA 1435
| | | | |
QY 1127 AGATATTAAGCTATTGAAAACTAGTATCGGTGACACATCAATCTCAAGAAATTA 1186
| | | | |
Db 1436 AGATATTAAGCTATTGAAAACTAGTATCGGTGACACATCAATCTCAAGAAATTA 1495
| | | | |
QY 1187 AGCTCAAGACAAAGATTTTAAACAACCAACCCGCTATACGATTTATGAAAGTGA 1246
| | | | |
Db 1496 AGCTCAAGACAAAGATTTTAAACAACCAACCCGCTATACGATTTATGAAAGTGA 1555
| | | | |
QY 1247 CTCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGATCAAGA 1306
| | | | |
Db 1556 CTCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGATCAAGA 1615
| | | | |
QY 1307 GTTTACTTACCGTGTAAAAATCGGAAACAAGCTTATAGATCAATAAAAATCTGTCT 1366
| | | | |
Db 1616 GTTTACTTACCGTGTAAAAATCGGAAACAAGCTTATAGATCAATAAAAATCTGTCT 1675
| | | | |
QY 1367 GAATGAAGAAATTAACAACACTGACCTGATCTGAGAAATATTAAGTCTTTAAAAAGG 1426
| | | | |
Db 1676 GAATGAAGAAATTAACAACACTGACCTGATCTGAGAAATATTAAGTCTTTAAAAAGG 1735
| | | | |
QY 1427 GGAAGAGCGTATGATCCCTTGTATGATGACGATCTTGAACCTGTTCAACCAATATAGT 1486
| | | | |
Db 1736 GGAAGAGCGTATGATCCCTTGTATGATGACGATCTTGAACCTGTTCAACCAATATAGT 1795
| | | | |
QY 1487 TGAATGATACCAAGAAATTTGCTAAAAAGTGAAGAGCTTAAACAGTACGAACTTAA 1546
| | | | |

```

Db      1796  TGAATGCGATACCAAGAAATGCTAAAGTAGAGCACTTAAACAGTAGCAAGCTAA 1855
Oy      1547  CTTAGACTCAGAGATTTATAGCATCTCTGTGTAAGGCTAAACTCTACCAACATCT 1606
Db      1856  CTTAGACTCAGAGATTTATAGCATCTCTGTGTAAGGCTAAACTCTACCAACATCT 1915
Oy      1607  CGATGCTTTGGTATATAGGACTATACCTTAACCTGAAAAAGTAGAGATATACAGATGA 1666
Db      1916  CGATGCTTTGGTATATAGGACTATACCTTAACCTGAAAAAGTAGAGATATACAGATGA 1975
Oy      1667  CACCAACCGTATCATACCGTTATATGCGCAAGCAGCCGGAAGAGAGAAATGCTAGCTA 1726
Db      1976  CACCAACCGTATCATACCGTTATATGCGCAAGCAGCCGGAAGAGAGAAATGCTAGCTA 2035
Oy      1727  CCATTTAGCTGTGCT 1742
Db      2036  TCATTTAGCTTATGAT 2051

RESULT 4
SEDEXB  8931 bp      DNA      linear      BCT 18-APR-2005
LOCUS   S.equisimilis dexb, abc, lrp, skc, rel genes and ORF1.
DEFINITION
ACCESSION X72832
VERSION    X72832.1 GI:407876
KEYWORDS   abc gene; ABC transporter; dexB gene; dextran glucosidase; glucan
            1,6-alpha-glucosidase; leucine rich protein; LRP gene; ORF1; rel
            gene; skc gene; streptokinase; stringent response-like protein.
SOURCE     Streptococcus dysgalactiae subsp. equisimilis
ORGANISM   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
            Streptococcus.
REFERENCE 1 (bases 3621 to 6190)
AUTHORS   Malke,H., Roe,B. and Ferretti,J.J.
TITLE     Nucleotide sequence of the streptokinase gene from Streptococcus
            equisimilis H46A
JOURNAL   Gene 34 (2-3), 357-362 (1985)
PUBMED   2989113
REFERENCE 2 (bases 1 to 4188, 5790 to 8931)
AUTHORS   Mechohl,U., Steiner,K., Vettermann,S. and Malke,H.
TITLE     Genetic organization of the streptokinase region of the
            Streptococcus equisimilis H46A chromosome
JOURNAL   Mol.Gen. Genet. 241 (1-2), 129-140 (1993)
PUBMED   8232196
REFERENCE 3 (bases 1 to 8931)
AUTHORS   Malke,H.
TITLE     Direct Submission
JOURNAL   Submitted (05-MAY-1993) H. Malke, Institute for Molecular Biology,
            Jena University, Winzerlaer Str 10, 07708 Jena, FRG
COMMENT    Related sequences: K02986, M19346, X13399 & X13400.
FEATURES   Location/Qualifiers
            1..8931
                /organism="Streptococcus dysgalactiae subsp. equisimilis"
                /mol_type="genomic DNA"
                /strain="H46A"
                /isolate="human group C strain"
                /sub_species="equisimilis"
                /db_xref="taxon:119602"
                /chromosome="streptokinase region"
                /clone_lib="E.coli plasmid library containing subfragments
                of the submitted seq:PSHD14/16, PSPV19, PSR2, PMR1,
                PCWP73, PRH10, PMX4"
                /complement(89..1761)
                /gene="dexB"
                /complement(89..115)
                /gene="dexB"
                /complement(89..115)
                /note="hairpin loop"
                /complement(136..1749)
                /gene="dexB"
                /BC_number="3.2.1.70"
                /codon_start=1
                /transl_table=11

            misc_feature
                /product="glucan 1,6-alpha-glucosidase"
                /protein_id="CA51348.1"
                /db_xref="GI:407877"
                /db_xref="GOA:Q59905"
                /db_xref="InterPro:IPR006047"
                /db_xref="InterPro:IPR006589"
                /db_xref="UniProt/Swiss-Prot:O59905"
                /translation="MOKOMHAKATVQIYPRSPKDTSGNGIGDLKGTISOLDYLOKLG
                ITAIVLSPVQSPMDNGYDIDSEYALAEVPGNMDMDLLAANERGKIKTIMDLVYN
                HTSDEHAMFVEABENPNSPRDYITRDEPNMLMSIFSGAMELDEASGOYIHLFSK
                KQPDILWENAHVRQKJYDWMNFWIAGIGFPRMDVIDLICKIPDSEITGNGPRHLDTL
                KENNQATPFGNHDVMTVGETGATPEIARQYRPNKELSMVFQFEHVGLQHPKPAKPM
                DYAEQDLVPALKITIFSKMOTELKLGEMNLSLFWNNHDLPRVLSIWMGDSIYRESGAQA
                LAILIHMRCGPYIYQGEELIGNTNYPKDLTEVDLIESLVAKEAMNGVPAALVMS
                IRRVGRDNATPQMGSKDTHTAGSEAEETLPVNPYQELINVADALANQSIPTTYQO
                LIAKRDQDMLEADYHLPTADKVRVYQRFGEETIYIVNVSDQVQPAKDLAAG
                VVITNTDVKVLETKHLQPMDAFCVKLSV"
                /complement(1757..1761)
                /gene="dexB"
                /complement(1780..3051)
                /gene="abc"
                /complement(1780..1810)
                /gene="abc"
                /note="hairpin loop"
                /complement(2482..2499)
                /gene="abc"
                /note="Walker motif B"
                /complement(2633..2965)
                /gene="abc"
                /codon_start=1
                /transl_table=11
                /product="ABC transporter"
                /protein_id="CA51349.1"
                /db_xref="GI:600072"
                /db_xref="GOA:Q54086"
                /db_xref="InterPro:IPR003439"
                /db_xref="UniProt/TREMBL:O54086"
                /translation="MVEINLNHIYKYPNTTHYAVEDFDIDKKEIVVPGSGCK
                STLRMTAGLEDISEGLKIDGVSVDKSPKDIAMFQNTALYPMHTYDNNAPGL
                KLRKYKR"
                /complement(2644)
                /gene="abc"
                /note="(+) frame shift mutation in H46A"
                /complement(2828..2854)
                /gene="abc"
                /note="Walker motif A"
                /complement(2973..2976)
                /gene="abc"
                /complement(3026..3031)
                /gene="abc"
                /note="alternative"
                /complement(3043..3048)
                /gene="abc"
                /complement(3050..3051)
                /gene="abc"
                /note="TG motif"
                /complement(3244..4149)
                /gene="lrp"
                /complement(3244..3261)
                /gene="lrp"
                /note="hairpin loop"
                /complement(3262..4110)
                /gene="lrp"
                /codon_start=1
                /transl_table=11
                /product="leucine rich protein"
                /protein_id="CA51350.1"
                /db_xref="GI:407878"
                /db_xref="GOA:Q54087"
                /db_xref="InterPro:IPR002197"
                /db_xref="InterPro:IPR008931"
                /db_xref="UniProt/Swiss-Prot:O54087"
                /translation="MELKDYFPEMQNGPHLGDKEWVSVEKGDQYVHPKSCLSKKR"

```

misc_feature	complement(3337. .3422) /gene="ltp" /note="leucine zipper motif"
RBS	complement(411. .4120) /gene="ltp" complement(4141. .4146) /gene="ltp" complement(4148. .4149) /note="ltp" /gene="ltp" /note="TG motif"
misc_feature	4172. .4368 /note="static DNA bending locus"
misc_feature	4392. .5837 /gene="akc" 4392. .4393 /gene="akc" /note="TG motif"
misc_feature	4395. .4400 /gene="akc" 4401. .4406 /gene="akc" /note="alternative"
-10_signal	4428. .4431 /gene="akc"
-10_signal	4439. .5761 /gene="akc" /codon_start=1 /transl_table=1 /product="streptokinase" /protein_id="CAA51351.1" /db_xref="GI:407879" /db_xref="GOA:P00779" /db_xref="InterPro:IPR004093" /db_xref="InterPro:IPR008124" /db_xref="UniProt/Swiss-Prot:P00779"
RBS	translation="MKVYLSFGMFLPALPFGTVNSVOALAGPEWLLDRSPVNSQULVSVAGYBGTGNODISLKFPFIDLTSRPAHGKTBQGLSPSKFPATDSGMSHLEKADLVAITLBQULANVHNSDDYFEVIDTFASDATITDRNGKVFADKDSVYLPTPVDEFLLSHVRFRPRPEKPIQNOAKSVDEYVQPTPLNPDDPRPGIKOTKLKTLAIGDITTSQELAAQOISLNKNHPGTYTIERDSSIVTHNDNIFRTILPMDEQFTYRVNRCQOAYRIKQASQIENINNTDLSIEKYVLLKKGEKPYDPDRSHLKLFTIKYVDNTEILLYSQELLTTSERNLDPRDLYDPRDKLLNNLDAGINDTYLTGKVEDNHDTRIRITTYMGKRPGEANSTHAIYDKDKYTEEREVSYLNTYTFPIPNPNDR"
sig_peptide	4439. .4516 /gene="akc"
terminator	5796. .5837 /gene="akc"
CDS	complement(5859. .6302) /note="unnamed protein product; ORF1" /codon_start=1 /transl_table=1 /protein_id="CAA51352.1" /db_xref="GI:407880" /db_xref="GOA:O54088" /db_xref="InterPro:IPR003732" /db_xref="UniProt/Swiss-Prot:O54088" translation="MKVLQRYKKEVSIIDKIGAIINQGLLVGVGDDPAADLAVYAVRKLVNKRISDADGKNQOSIOIKGSIISVSQPTLYADTKGNRPAPFGAANPDMA SQFYRPFNEQLADPVPYRAGVFGADMQVSLINDGVTIILDTKH"
conflict	6007. .6009 /citation=[1] /replace="gc"
conflict	6162. .6164 /citation=[1] /replace="cg"
RBS	complement(6310. .6313) complement(6317. .8575) /note="ORF1"
gene	

Query Match	54.7%	Score 1146.4	DB 1	Length 8911	
Best Local Similarity	99.5%	Pred. No. 1.3e-297			
Matches 1150	Conservative	0	Mismatches 6	Indels 0	Gaps 0
QY 587 TATTCGTGACCTGATGAGTGGCTGCTGAGACCGTGCATCTGTCACAAACAGCCATTTGGTGT					646
Db 4516 TATTCGTGACCTGATGAGTGGCTGCTGAGACCGTGCATCTGTCACAAACAGCCATTTAGTGT					4575
QY 647 TAGCGTGTCTGATCTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGGAAAT					706
Db 4576 TAGCGTGTCTGATCTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGGAAAT					4639
QY 707 CGATCTAACATACGACCTGCTCATATGAGGAAAGACAGGCAAGGCTTAATGTCGAAATC					766
Db 4636 CGATCTAACATACGACCTGCTCATATGAGGAAAGACAGGCAAGGCTTAATGTCGAAATC					4699
QY 767 AAAACCATTTGCTACTGATATAGTGGCGCGATGTCATATACTTGAAGAAAGCTGACTTACT					826
Db 4696 AAAACCATTTGCTACTGATATAGTGGCGCGATGTCATATACTTGAAGAAAGCTGACTTACT					4755
QY 827 AAAGGCTATTCAGAAACAATTGATCGTAAACGTCCACAGTAAACGACATCTATTGAGGT					886
Db 4756 AAAGGCTATTCAGAAACAATTGATCGTAAACGTCCACAGTAAACGACATCTATTGAGGT					4815
QY 887 CATTGATTTTGCACCGATGCAACCATTACTGATGAAACGGCAAGTCTTACTTTGCTGCA					946
Db 4816 CATTGATTTTGCACCGATGCAACCATTACTGATGAAACGGCAAGTCTTACTTTGCTGCA					4875
QY 947 CAAAGATGGTGGGTAAACCTTGGCGACCAACCTGTCGAAAGAAATTTTGTCTAAGCGGACA					1006
Db 4876 CAAAGATGGTGGGTAAACCTTGGCGACCAACCTGTCGAAAGAAATTTTGTCTAAGCGGACA					4935
QY 1007 TGTGGCGCTTAGACCATATTAAGAAACCAATACAAACCAAGCGAAATCTGTGATGT					1066
Db 4936 TGTGGCGCTTAGACCATATTAAGAAACCAATACAAACCAAGCGAAATCTGTGATGT					4995
QY 1067 GGAATATATCTGTACAGTTTACTCCCTTAAACCTGTGATGACGATTTGACACAGGCTGCA					1126
Db 4996 GGAATATATCTGTACAGTTTACTCCCTTAAACCTGTGATGACGATTTGACACAGGCTGCA					5055
QY 1127 AGATCTAAGCAATTAAGAAACCACTAGCTATGGGTGACACATCAATCTCAAGAAATTACT					1186
Db 5056 AGATCTAAGCAATTAAGAAACCACTAGCTATGGGTGACACATCAATCTCAAGAAATTACT					5115
QY 1187 AGCTCAAGCAACAAGACATTTTAAACAAACCAACCCAGGCTTAACGATTTATGACGTA					1246
Db 5116 AGCTCAAGCAACAAGACATTTTAAACAAACCAACCCAGGCTTAACGATTTATGACGTA					5175
QY 1247 CTCTCTCAATCTGCTACTCATGCAATGACATTTTCCGTACGATTCGATTCGATTCGATTCGAA					1306
Db 5176 CTCTCTCAATCTGCTACTCATGCAATGACATTTTCCGTACGATTCGATTCGATTCGATTCGAA					5235
QY 1307 GTTTACTTAACCGTGTAAATAATCGGGGACAAAGCTTATAGATTCATATAAAATCTGCTCT					1366
Db 5236 GTTTACTTAACCGTGTAAATAATCGGGGACAAAGCTTATAGATTCATATAAAATCTGCTCT					5295
QY 1367 GAATGAAGAAATTAACCAACACTGACCTGATCTCTGAGAAATTAATGCTCTTAAAAAAG					1426
Db 5296 GAATGAAGAAATTAACCAACACTGACCTGATCTCTGAGAAATTAATGCTCTTAAAAAAG					5355
QY 1427 GGAAGAGCCGTATGATCCCTTTGATCGAGTCACTTGAATCTGTACCATCAATACATACGT					1486
Db 5356 GGAAGAGCCGTATGATCCCTTTGATCGAGTCACTTGAATCTGTACCATCAATACATACGT					5415
QY 1487 TGAATGCAATACCAACGAATTCGCTTAAAAAGTGAAGAGAGCTCTTAAACAGTACGGAACGTAA					1546
Db 5416 TGAATGCAATACCAACGAATTCGCTTAAAAAGTGAAGAGAGAGCTCTTAAACAGTACGGAACGTAA					5475
QY 1547 CTTGAAGCTTCAAGATTTATACGATCTCGTATGAAGGCTTAAAGCTACTCTTCAACAATCT					1606

```

Db      5476 CTTAGACTTCAGAGATTATACGATCTCGTGATTAAGGCTAAACTACTCTACCAACAATCT 5535
Qy      1607 CGATGCTTTGGATTATAGACATATACCTTAACCTGAGAAAGTAGAGATATACCGATGA 1666
Db      5536 CGATGCTTTGGATTATAGACATATACCTTAACCTGAGAAAGTAGAGATATACCGATGA 5595
Qy      1667 CACCAACCGTATCATATACGTTATATAGGCAAGCCGCAAGAGAGAGATGCTACTA 1726
Db      5596 CACCAACCGTATCATATACGTTATATAGGCAAGCCGCAAGAGAGAGATGCTACTA 5655
Qy      1727 CCATTTAGCTGTGT 1742
Db      5656 TCATTTAGCCTATGAT 5671

RESULT 5
ARI44000
LOCUS      ARI44000      1242 bp      DNA      linear      PAT 08-AUG-2001
DEFINITION Sequence 5 from patent US 6210667.
ACCESSION ARI44000
VERSION    ARI44000.1 GI:15105867
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unknown.
REFERENCE    1 (bases 1 to 1242)
              Unclassified.
AUTHORS      Reed,G.L.
TITLE        Bacterial fibrin-dependent plasminogen activator
JOURNAL      Patent: US 6210667-A 5 03-APR-2001;
FEATURES
            source      1..1242
                        /organism="Unknown"
                        /mol_type="unassigned DNA"

Query Match      54.6%; Score 1145.4; DB 6; Length 1242;
Best Local Similarity 99.5%; Pred. No. 2.1e-297;
Matches 1149; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy      588 ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAAATGTTGT 647
Db      1 ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAAATGTTGT 60
Qy      648 ACCGTTGCTGTACTGTGAGGGGAGAGATCAAGACATTTGCTTAAATTTTGAATC 707
Db      61 ACCGTTGCTGTACTGTGAGGGGAGAGATCAAGACATTTGCTTAAATTTTGAATC 120
Qy      708 GATCTAACATCAAGACCTGCTCATGAGAGAAAGACAGAGAGGCTTAAGTCCAAATCA 767
Db      121 GATCTAACATCAAGACCTGCTCATGAGAGAAAGACAGAGAGGCTTAAGTCCAAATCA 180
Qy      768 AAACCATTTGCTACTGATAGTGGCGGAGATGTCACATTAACCTTGAGAAAGCTGACTTACTA 827
Db      181 AAACCATTTGCTACTGATAGTGGCGGAGATGTCACATTAACCTTGAGAAAGCTGACTTACTA 240
Qy      828 AAGGCTATTCAAGAACATTTGATGCTTAAGTCCACAGTACAGACGTACTTTAGAGTTC 887
Db      241 AAGGCTATTCAAGAACATTTGATGCTTAAGTCCACAGTACAGACGTACTTTAGAGTTC 300
Qy      888 ATTGATTTTGAAGGAGATGAACATTAAGTCAATGGAAGGCAAGGCTTACTTTGCTGAC 947
Db      301 ATTGATTTTGAAGGAGATGAACATTAAGTCAATGGAAGGCAAGGCTTACTTTGCTGAC 360
Qy      948 AAAGATGTTGGTGAACCTTGGCCAGCCAACTGTCCAAAGATTTTGTCTAAGGAGCAT 1007
Db      361 AAAGATGTTGGTGAACCTTGGCCAGCCAACTGTCCAAAGATTTTGTCTAAGGAGCAT 420
Qy      1008 GTGCGCGTTAGACCATATATAAGAAAAACAATACAAAACCAAGCGAAATCTGTTGATGTG 1067
Db      421 GTGCGCGTTAGACCATATATAAGAAAAACAATACAAAACCAAGCGAAATCTGTTGATGTG 480
Qy      1068 GAATATACGTATACGTTTACTCCCTTAAACCTGATGACGATTTGACACAGGCTCTCAA 1127

```

```

Db      481 GAATATACGTATACGTTTACTCCCTTAAACCTGATGACGATTTGACACGAGTCTCAA 540
Qy      1128 GATCTAGCATATGAAAAACATGATATGCGTGAACCAACATCATCAAGATTTACTA 1187
Db      541 GATCTAGCATATGAAAAACATGATATGCGTGAACCAACATCATCAAGATTTACTA 600
Qy      1188 GCTCAAGCAACAAGCATTTTAAACAAAAACCAAGGCTTATGAGATTTATGAAAGTAC 1247
Db      601 GCTCAAGCAACAAGCATTTTAAACAAAAACCAAGGCTTATGAGATTTATGAAAGTAC 660
Qy      1248 TCCTCAATCGTCACTGACATGACATGACATTTCCGTACGATTTTACCAATGATCAAG 1307
Db      661 TCCTCAATCGTCACTGACATGACATGACATTTCCGTACGATTTTACCAATGATCAAG 720
Qy      1308 TTTACTTACCGGTGTTAAAAATCGGGAACAGGCTTATGAGATCAATTAATAAATCTGCTG 1367
Db      721 TTTACTTACCGGTGTTAAAAATCGGGAACAGGCTTATGAGATCAATTAATAAATCTGCTG 780
Qy      1368 AATGAGAAATTAACCAACACTGACCTGATCTGAGAAATATTAAGTCTTAAAAAAGG 1427
Db      781 AATGAGAAATTAACCAACACTGACCTGATCTGAGAAATATTAAGTCTTAAAAAAGG 840
Qy      1428 GAAAAAGCGTATGATCCCTTTGATGCGAGTCACTTGAACCTTTCAACCATCAATACGTT 1487
Db      841 GAAAAAGCGTATGATCCCTTTGATGCGAGTCACTTGAACCTTTCAACCATCAATACGTT 900
Qy      1488 GATGTCATACCAAGCAATGCTTAAAAAGTAGAGAGCTTTAAACGCTAGGCAAGCTAAC 1547
Db      901 GATGTCATACCAAGCAATGCTTAAAAAGTAGAGAGCTTTAAACGCTAGGCAAGCTAAC 960
Qy      1548 TTAGACTTCAGAGATTATACGATCTCGTGAATAGGCTTAAGCTTACATCAACAATCTC 1607
Db      961 TTAGACTTCAGAGATTATACGATCTCGTGAATAGGCTTAAAGCTTACATCAACAATCTC 1020
Qy      1608 GATGCTTTGGTATATAGACTATACCTTAACCTGAAAAAGTAGAGATATACAGATGAC 1667
Db      1021 GATGCTTTGGTATATAGACTATACCTTAACCTGAAAAAGTAGAGATATACAGATGAC 1080
Qy      1668 ACCAACCGTATCATTAACCGTTTATATAGGCAAGCGACCCGGAAGAGAGATGCTAGTAC 1727
Db      1081 ACCAACCGTATCATTAACCGTTTATATAGGCAAGCGACCCGGAAGAGAGATGCTAGTAT 1140
Qy      1728 CATTAGCTGTGT 1742
Db      1141 CATTAGCTGTATGAT 1155

RESULT 6
E00522
LOCUS      E00522      2568 bp      DNA      linear      PAT 29-SEP-1997
DEFINITION DNA fragment of plasmid pMF1 into which DNA encoding streptokinase
              is inserted.
ACCESSION  E00522
VERSION    E00522.1 GI:2168801
KEYWORDS   UP 1985237995-A/1.
SOURCE      Streptococcus dysgalactiae subsp. equisimilis
ORGANISM    Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
              Streptococcus.
REFERENCE   1 (bases 1 to 2568)
AUTHORS     Jiyosefu,J.P. and Horusuto,M.
TITLE       RECOMBINANT VECTOR FOR PRODUCING STREPTOKINASE
JOURNAL     PHILIPS PETROLEUM CO
COMMENT     OS Streptococcus equisimilis
              PN JP 1985237995-A/1
              PD 26-NOV-1985
              PF 09-OCT-1984 JP 1984212403
              PR 10-OCT-1983 DD 83 255523, 02-MAR-1984 US 84 585417 PI
              JIYOSSEFU JIEI FUJIRETSUTERI, HORUSUTO MARUKE
              PC C12N15/00,C12N1/20,C12N9/70,(C12N15/00,C12R1.46),(C12N1/20,PC
              C12R1.19),

```

PC (C12N9/70,C12R1:19);
 CC strandedness: Single;
 CC topology: Linear;
 CC hypothetical: No;
 CC anti-sense: No;
 CC *source: Strain=Streptococcus equismilis H46A; CC *source: library=Streptococcus equismilis H46 library; CC *source: clone=Lambda L47 skc clone;
 FH Key Location/Qualifiers
 FT 5'UTR 1..818
 FT sig_peptide 819..896
 FT /product='streptokinase signal peptide' FT
 mat_peptide 897..2138
 FT /product='streptokinase'
 FT CDS 819..2143
 FT /product='streptokinase'
 FT 3'UTR 2142..2569
 Location/Qualifiers
 1..2568
 /organism='Streptococcus dysgalactiae subsp. equismilis'
 /mol_type='genomic DNA'
 /db_xref='taxon:119602'

ORIGIN

Query Match 54.5%; Score 1141.6; DB 6; Length 2568;
 Best Local Similarity 99.2%; Pred. No. 2.3e-295;
 Matches 1147; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

587 TATGTGAGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCAATTTGTTGT 646
 DB TATGTGAGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCAATTTGTTGT 955
 647 TAGCGTGTGCTGATCTGTGAGGGAGCAATCAAGACATTAGTCTTAATTTTGAAT 706
 DB TAGCGTGTGCTGATCTGTGAGGGAGCAATCAAGACATTAGTCTTAATTTTGAAT 1015
 956 TAGCGTGTGCTGATCTGTGAGGGAGCAATCAAGACATTAGTCTTAATTTTGAAT 1015
 707 CGATCTAATCATCAAGACCTGCTCATGGAGGAAGAGCAAGGCTTAGTCCAAATC 766
 DB CGATCTAATCATCAAGACCTGCTCATGGAGGAAGAGCAAGGCTTAGTCCAAATC 1075
 1016 CGATCTAATCATCAAGACCTGCTCATGGAGGAAGAGCAAGGCTTAGTCCAAATC 1075
 767 AAAACATTGTGCTACTGATAGTGGCCGATGTCAATTAACCTTGAGAACTGACTTACT 826
 DB AAAACATTGTGCTACTGATAGTGGCCGATGTCAATTAACCTTGAGAACTGACTTACT 1135
 1076 AAAACATTGTGCTACTGATAGTGGCCGATGTCAATTAACCTTGAGAACTGACTTACT 1135
 827 AAAAGCTATTCAAGAACATTTGCTTAACTGCTCAAGTCAAGTCAAGTCAAGTCAAGT 886
 DB AAAAGCTATTCAAGAACATTTGCTTAACTGCTCAAGTCAAGTCAAGTCAAGTCAAGT 1195
 1136 AAAAGCTATTCAAGAACATTTGCTTAACTGCTCAAGTCAAGTCAAGTCAAGTCAAGT 1195
 887 CATTGATTTTGCAGAGGATGCAACATTAATCTGATTCGAAAGCGAAGTCTTACTTGTGA 946
 DB CATTGATTTTGCAGAGGATGCAACATTAATCTGATTCGAAAGCGAAGTCTTACTTGTGA 1255
 1196 CATTGATTTTGCAGAGGATGCAACATTAATCTGATTCGAAAGCGAAGTCTTACTTGTGA 1255
 947 CAAAGTGTGCTGCTGACCTGCGGACCCCAACCTGTCAGAAATTTTGTCTAAGCGGACA 1006
 DB CAAAGTGTGCTGCTGACCTGCGGACCCCAACCTGTCAGAAATTTTGTCTAAGCGGACA 1315
 1256 CAAAGTGTGCTGCTGACCTGCGGACCCCAACCTGTCAGAAATTTTGTCTAAGCGGACA 1315
 1007 TGTGGCGCTTGAACCATATTAAGAAAAACCATATCAAAACCAAGCAATCTGTTGATGT 1066
 DB TGTGGCGCTTGAACCATATTAAGAAAAACCATATCAAAACCAAGCAATCTGTTGATGT 1375
 1316 TGTGGCGCTTGAACCATATTAAGAAAAACCATATCAAAACCAAGCAATCTGTTGATGT 1375
 1067 GGAATATATCTGTAAGTTTACTCCCTTAAACCTGATGACGATTTTGAACAGGCTTCAA 1126
 DB GGAATATATCTGTAAGTTTACTCCCTTAAACCTGATGACGATTTTGAACAGGCTTCAA 1435
 1376 GGAATATATCTGTAAGTTTACTCCCTTAAACCTGATGACGATTTTGAACAGGCTTCAA 1435
 1127 AGATCTAGATCTATTAAGAAACATAGCTATCGGTGACACATGACATCTTAAAGTAACT 1186
 DB AGATCTAGATCTATTAAGAAACATAGCTATCGGTGACACATGACATCTTAAAGTAACT 1495
 1436 AGATCTAGATCTATTAAGAAACATAGCTATCGGTGACACATGACATCTTAAAGTAACT 1495
 1187 AGCTCAAGCAAAAGATTTTAAACAAACCAAGGCTATACGATTTATGAAACGTGA 1246
 DB AGCTCAAGCAAAAGATTTTAAACAAACCAAGGCTATACGATTTATGAAACGTGA 1555
 1496 AGCTCAAGCAAAAGATTTTAAACAAACCAAGGCTATACGATTTATGAAACGTGA 1555

1247 CTCCTCAATCTGCTACATGACAAATGATTTTCGATGATTTTACCAATGATCAAGA 1306
 DB CTCCTCAATCTGCTACATGACAAATGATTTTCGATGATTTTACCAATGATCAAGA 1615
 1556 CTCCTCAATCTGCTACATGACAAATGATTTTCGATGATTTTACCAATGATCAAGA 1615
 1307 GTTTACTTACCGGTGTTAAATTCGGAAACAAGCTTATGATCAATTAATAATCTGTCT 1366
 DB GTTTACTTACCGGTGTTAAATTCGGAAACAAGCTTATGATCAATTAATAATCTGTCT 1675
 1616 GTTTACTTACCGGTGTTAAATTCGGAAACAAGCTTATGATCAATTAATAATCTGTCT 1675
 1367 GAATGAAGAAATTAACAAACATGACCTGATCTCTGAGAAATTAATGCTTAAATAAG 1426
 DB GAATGAAGAAATTAACAAACATGACCTGATCTCTGAGAAATTAATGCTTAAATAAG 1735
 1676 GAATGAAGAAATTAACAAACATGACCTGATCTCTGAGAAATTAATGCTTAAATAAG 1735
 1427 GGAAGACCGTATGATCCCTTTGATTCGAGTCACTTGAACGTTCACCATCAATAGCT 1486
 DB GGAAGACCGTATGATCCCTTTGATTCGAGTCACTTGAACGTTCACCATCAATAGCT 1795
 1736 GGAAGACCGTATGATCCCTTTGATTCGAGTCACTTGAACGTTCACCATCAATAGCT 1795
 1487 TGATGTGATTCACCAAGATTTGCTTAAAGTGAAGAGCTTTTAAAGCTTACGAAAGCT 1546
 DB TGATGTGATTCACCAAGATTTGCTTAAAGTGAAGAGCTTTTAAAGCTTACGAAAGCT 1855
 1796 TGATGTGATTCACCAAGATTTGCTTAAAGTGAAGAGCTTTTAAAGCTTACGAAAGCT 1855
 1547 CTTGACTTTCAGAGATTTATTAAGATCCCTGATGATTAAGGCTTAACTTCTCAACATCT 1606
 DB CTTGACTTTCAGAGATTTATTAAGATCCCTGATGATTAAGGCTTAACTTCTCAACATCT 1915
 1856 CTTGACTTTCAGAGATTTATTAAGATCCCTGATGATTAAGGCTTAACTTCTCAACATCT 1915
 1607 CGATGCTTTTGGTATTTATGACATTAACCTTAATGAGAAAGTAAAGATTAATCAAGTGA 1666
 DB CGATGCTTTTGGTATTTATGACATTAACCTTAATGAGAAAGTAAAGATTAATCAAGTGA 1975
 1916 CGATGCTTTTGGTATTTATGACATTAACCTTAATGAGAAAGTAAAGATTAATCAAGTGA 1975
 1667 CACCAACCGTATCAATTAACCTTTTATATGAGGAGAGGACCCGAGAGAGAAATGCTAGCTA 1726
 DB CACCAACCGTATCAATTAACCTTTTATATGAGGAGAGGACCCGAGAGAGAAATGCTAGCTA 2035
 1976 CACCAACCGTATCAATTAACCTTTTATATGAGGAGAGGACCCGAGAGAGAAATGCTAGCTA 2035
 1727 CCATTTAGCTGTGTGT 1742
 DB CCATTTAGCTGTGTGT 2051

RESULT 7
 LOCUS 105204 1401 bp DNA linear PAT 02-DEC-1994
 DEFINITION Sequence 5 from Patent EP 0248227.
 ACCESSION 105204
 VERSION 105204.1 GI:591209
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1401)
 AUTHORS Hagenstrom, M.J. and Stromman, D.W.
 TITLE Yeast production of streptokinase
 JOURNAL Patent: EP 0248227-A1 5 09-DEC-1987;
 FEATURES
 source 1..1401
 /organism='unknown'
 /mol_type='unassigned DNA'

ORIGIN

Query Match 54.3%; Score 1139; DB 6; Length 1401;
 Best Local Similarity 99.1%; Pred. No. 1.1e-295;
 Matches 1145; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

588 ATTGCTGACCTGAGTGGCTGTGACCGTCCATCTGTCAACAAAGCAATGTTGTT 647
 DB ATTGCTGACCTGAGTGGCTGTGACCGTCCATCTGTCAACAAAGCAATGTTGTT 69
 10 ATTGCTGACCTGAGTGGCTGTGACCGTCCATCTGTCAACAAAGCAATGTTGTT 69
 648 AGCTTCTGTGATCTGTTGAGGGGAGCAATTAAGACATTAATTTTGAATC 707
 DB AGCTTCTGTGATCTGTTGAGGGGAGCAATTAAGACATTAATTTTGAATC 129
 70 AGCTTCTGTGATCTGTTGAGGGGAGCAATTAAGACATTAATTTTGAATC 129
 708 GATTAACATCAAGACCTGCTCATGAGAGAAAGCAAGCAAGGCTTAAGTCAAAATCA 767
 DB GATTAACATCAAGACCTGCTCATGAGAGAAAGCAAGCAAGGCTTAAGTCAAAATCA 189
 130 GATTAACATCAAGACCTGCTCATGAGAGAAAGCAAGCAAGGCTTAAGTCAAAATCA 189

QY 768 AAACATTTGCTACTGATAGTGGCGGATGTCACTATACTTGAGAAAGCTGACTTACTA 827
DB 190 AAACATTTGCTACTGATAGTGGCGGATGTCACTATACTTGAGAAAGCTGACTTACTA 249
QY 828 AAGGCTATTGAGAAACAATTGATCGCTAACGTCACAGTACGACTTACCTTTGAGGTC 887
DB 250 AAGGCTATTGAGAAACAATTGATCGCTAACGTCACAGTACGACTTACCTTTGAGGTC 309
QY 888 ATTGATTTTGGAGGAGTCAACCACTTACGATGGAAGCGGAGGCTTACTTTGCTGAC 947
DB 310 ATTGATTTTGGAGGAGTCAACCACTTACGATGGAAGCGGAGGCTTACTTTGCTGAC 369
QY 948 AAAGATGGTTCGGTAACTTTCGCGAACCACTGTCACAGAAATTTTGGCTAACGCGACAT 1007
DB 370 AAAGATGGTTCGGTAACTTTCGCGAACCACTGTCACAGAAATTTTGGCTAACGCGACAT 429
QY 1008 GTGGCGGTTGAGCCATATTAAGAAAAACAATACAAAACCAAGCGAAATCTGTTGATGTG 1067
DB 430 GTGGCGGTTGAGCCATATTAAGAAAAACAATACAAAACCAAGCGAAATCTGTTGATGTG 489
QY 1068 GAATATTAAGCTATGTTAACTCCCTTAAACCTTGATGACGATTTTCAACCGGCTTCAAA 1127
DB 490 GAATATTAAGCTATGTTAACTCCCTTAAACCTTGATGACGATTTTCAACCGGCTTCAAA 549
QY 1128 GATATTAAGCTATGTTAACTAGGCTATCGGTGACACCATCACTCTCAAGAAATTAATA 1187
DB 550 GATATTAAGCTATGTTAACTAGGCTATCGGTGACACCATCACTCTCAAGAAATTAATA 609
QY 1188 GCTCAAGCACAAGCAATTTTAAACAAACCAACCGGCTATAAGATTATGAACGTGAC 1247
DB 610 GCTCAAGCACAAGCAATTTTAAACAAACCAACCGGCTATAAGATTATGAACGTGAC 669
QY 1248 TCCTCAATGCTATCATGACATGACATTTTCCGTAACATTTTACCAATGATCAAGAG 1307
DB 670 TCCTCAATGCTATCATGACATGACATTTTCCGTAACATTTTACCAATGATCAAGAG 729
QY 1308 TTTACTTACCGTGTAAAAATCGGGAACAGCTTATAGATCAATTAATAAATCGGTG 1367
DB 730 TTTACTTACCGTGTAAAAATCGGGAACAGCTTATAGATCAATTAATAAATCGGTG 789
QY 1368 AATGAAGAAATTAACAACAACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1427
DB 790 AATGAAGAAATTAACAACAACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 849
QY 1428 GAAAAGCCGTATGATCCCTTTGATCGGAGTCACTTGAACCTGTCACCAATAATGCTT 1487
DB 850 GAAAAGCCGTATGATCCCTTTGATCGGAGTCACTTGAACCTGTCACCAATAATGCTT 909
QY 1488 GATGTCGATACCAAGAAATTTGCTAAAAAGTGAAGCAGCTCTTACAGCTGCGAAGCTAAC 1547
DB 910 GATGTCGATACCAAGAAATTTGCTAAAAAGTGAAGCAGCTCTTACAGCTGCGAAGCTAAC 969
QY 1548 TTAGACTTCAGAGATTTATACGATCCCTGATTAAGGCTTAACTACTCTCAACAATCTC 1607
DB 970 TTAGACTTCAGAGATTTATACGATCCCTGATTAAGGCTTAACTACTCTCAACAATCTC 1029
QY 1608 GATGCTTTTGTATTAATGACATATACCTTAACCTGAAAAAGTGAAGATTAATCAAGTAC 1667
DB 1030 GATGCTTTTGTATTAATGACATATACCTTAACCTGAAAAAGTGAAGATTAATCAAGTAC 1089
QY 1668 ACCAACCGTATACATAACCGTTTATATGGGCAAGCGAACCGGAAAGAAATGCTACTGAC 1727
DB 1090 ACCAACCGTATACATAACCGTTTATATGGGCAAGCGAACCGGAAAGAAATGCTACTGAC 1149
QY 1728 CATTTAGCTGTGCT 1742
DB 1150 CATTTAGCTGTGCT 1164

RESULT 8
AR068768 2566 bp DNA linear PAT 29-SEP-1999
LOCUS AR068768
DEFINITION Sequence 19 from patent US 5854049.

ACCESSION AR068768
VERSION AR068768.1 GI:6000975
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 2566)
TITLE Reel, G.L.
JOURNAL Plasmid-resistant streptokinase
PATENT Patent: US 5854049-A 19 29-DEC-1998;
FEATURES Location/Qualifiers
Source 1..2566
/organism="unknown"
/mol_type="unassigned DNA"

Query Match 54.0%; Score 1132.4; DB 6; Length 2566;
Best Local Similarity 99.3%; Pred. No. 7.1e-294;
Matches 1148; Conservative 0; Mismatches 6; Indels 2; Gaps 1;

QY 587 TATTGCTGACCTGAGTGGCTGCTGACCCGTCATCTGTCAACAACAGCCAAATGGTGT 646
DB 896 TATTGCTGACCTGAGTGGCTGCTGACCCGTCATCTGTCAACAACAGCCAAATGGTGT 955
QY 647 TAGGCTTCTGCTATGCTGTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAAT 706
DB 956 TAGGCTTCTGCTATGCTGTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAAT 1015
QY 707 GATCTAACATCAAGACCTGCTCAATGAGGAAAGACAGAGCAAGCTTAAGTCCAAATC 766
DB 1016 GATCTAACATCAAGACCTGCTCAT - AGGAAAGACAGAGCAAGCTTAAGTCCAAATC 1073
QY 767 AAAACATTTGCTATGATGCTGCGGAGTGTCACTATACTTGAAGAGCTTACT 826
DB 1074 AAAACATTTGCTATGATGCTGCGGAGTGTCACTATACTTGAAGAGCTTACT 1133
QY 827 AAAGGCTATTCAGAAACAATTGATGCTAAGCTCAAGTACGACGACTTGAAGT 886
DB 1134 AAAGGCTATTCAGAAACAATTGATGCTAAGCTCAAGTACGACGACTTGAAGT 1193
QY 887 CATTGATTTTGAAGAGCAGTCAACCATTAAGTCAAGACGGAAGGCTTACTTTGCTGA 946
DB 1194 CATTGATTTTGAAGAGCAGTCAACCATTAAGTCAAGACGGAAGGCTTACTTTGCTGA 1253
QY 947 CAAGATGTTTCGGTAACTTTCGCGAACCACTGTCTCAAGAAATTTTGTACGCGACA 1006
DB 1254 CAAGATGTTTCGGTAACTTTCGCGAACCACTGTCTCAAGAAATTTTGTACGCGACA 1313
QY 1007 TGTGCGGCTTAAGACCATATAAGAAACCAATTCAGAAACCAAGCGAAATCTGTGATGT 1066
DB 1314 TGTGCGGCTTAAGACCATATAAGAAACCAATTCAGAAACCAAGCGAAATCTGTGATGT 1373
QY 1067 GGAATATTAAGTACGATGTTTACTCCCTTAAACCTGATGACGATTTTGAAGAGCTTCAA 1126
DB 1374 GGAATATTAAGTACGATGTTTACTCCCTTAAACCTGATGACGATTTTGAAGAGCTTCAA 1433
QY 1127 AGATTAAGCTATTTGAAAAACCTAGCTATCGTGAACCAATTCAGATCTCAAGAAATCT 1186
DB 1434 AGATTAAGCTATTTGAAAAACCTAGCTATCGTGAACCAATTCAGATCTCAAGAAATCT 1493
QY 1187 AGCTCAAGGACAAAGCAATTTTAAACAAACCAACCGGCTATACGATTTTGAAGAGGGA 1246
DB 1494 AGCTCAAGGACAAAGCAATTTTAAACAAACCAACCGGCTATACGATTTTGAAGAGGGA 1553
QY 1247 CTCCTCAATGCTGACCTATGACCAATGACATTTTCCGTAAGATTTTACCAATGATCAAGA 1306
DB 1554 CTCCTCAATGCTGACCTATGACCAATGACATTTTCCGTAAGATTTTACCAATGATCAAGA 1613
QY 1307 GTTTACTTAACCGTGTAAAAATCGGGAACAAGCTTATAGATCAATATAAATATCTGCTCT 1366
DB 1614 GTTTACTTAACCGTGTAAAAATCGGGAACAAGCTTATAGATCAATATAAATATCTGCTCT 1673
QY 1367 GAATGAAGAAATTAACAACAACCTGATGCTGATCTCTGAGAAATATTACGTCCTTAAAAAGG 1426

Db 1674 GATATAGAAATTAACACACGACCTGATCTCTGAGAAATATTAGTCTTAAAAAAGG 1733
Qy 1427 GAAAAGCCGATGATCCCTTTGATCGACATGCACTTGAACCTGTCACATCAATACGT 1486
Db 1734 GAAAAGCCGATGATCCCTTTGATCGACATGCACTTGAACCTGTCACATCAATACGT 1793
Qy 1487 TGATGTCGATACCAAGAAATGCTAAAAAGTAGAGAGCTTTTAAACAGTAGCAAGTAA 1546
Db 1794 TGATGTCGATACCAAGAAATGCTAAAAAGTAGAGAGCTTTTAAACAGTAGCAAGTAA 1853
Qy 1547 CTTAAGCTTCAAGATTTTATAGATCTGCTGATAGAGCTTAACTTCTTCAACAATCT 1606
Db 1854 CTTAAGCTTCAAGATTTTATAGATCTGCTGATAGAGCTTAACTTCTTCAACAATCT 1913
Qy 1607 CGATCTTTTGGTATTTATGACCTTAACTTGAAGAGAGTAAATACAGATGA 1666
Db 1914 CGATCTTTTGGTATTTATGACCTTAACTTGAAGAGAGTAAATACAGATGA 1973
Qy 1667 CACCAACCGTATCATACCGTTTATATGGGCAAGCCGAGAGAGAGATCTAGCTA 1726
Db 1974 CACCAACCGTATCATACCGTTTATATGGGCAAGCCGAGAGAGAGATCTAGCTA 2033
Qy 1727 CCATTTAGCTGGTGT 1742
Db 2034 TCATTTAGCTATGAT 2049

RESULT 9

A20027 1458 bp DNA linear PAT 14-JUL-1995
LOCUS SEQ ID NO: 35; Synthetic nucleotide sequence for
DEFINITION Hirudin-streptokinase fusion protein.
A20027
ACCESSION A20027.1 GI:1247862
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 1458)
AUTHORS
TITLE
JOURNAL
FEATURES
SOURCE

ORIGIN

Query Match 53.7%; Score 1125.4; DB 6; Length 1458;
Best Local Similarity 97.8%; Pred. No. 5.2e-292;
Matches 1141; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
Qy 576 ACCGATGTTGCTGATGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGCAACAACAGC 635
Db 196 ATCGAAGGTAGATGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGCAACAACAGC 255
Qy 636 CAATTGCTTTAGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 695
Db 256 CAATTGCTTTAGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 315
Qy 696 TTTTGTGAATGATCTTAACATCAGACCTGCTCATGAGAGAGAGAGAGAGAGAGAGCTTA 755
Db 316 TTTTGTGAATGATCTTAACATCAGACCTGCTCATGAGAGAGAGAGAGAGAGAGAGCTTA 375
Qy 756 AGTCCAAATGAAAACCAATTTGCTACTGATAGTGGGCGCATGTCATATACTTGAAGAA 815
Db 376 AGTCCAAATGAAAACCAATTTGCTACTGATAGTGGGCGCATGTCATATACTTGAAGAA 435
Qy 816 GCTGACTTACTTAAGGCTATTGCAAGAACTATGATGCTTAAGCTTCAAGTAAAGAGAG 875
Db 436 GCTGACTTACTTAAGGCTATTGCAAGAACTATGATGCTTAAGCTTCAAGTAAAGAGAG 495

Qy 876 TACTTGAAGCTCATTTGATTTTGAAGAGATCAACCATTTACTGATCGAAAGGCAAGGTC 935
Db 496 TACTTGAAGCTCATTTGATTTTGAAGAGATCAACCATTTACTGATCGAAAGGCAAGGTC 555
Qy 936 TACTTGAAGCTCATTTGATTTTGAAGAGATCAACCATTTACTGATCGAAAGGCAAGGTC 995
Db 556 TACTTGAAGCTCATTTGATTTTGAAGAGATCAACCATTTACTGATCGAAAGGCAAGGTC 615
Qy 996 CTAAGCGGACATGCGCGGTAAACCATTTATGAAGAGAGAGAGAGAGAGAGAGAGAGAG 1055
Db 616 CTAAGCGGACATGCGCGGTAAACCATTTATGAAGAGAGAGAGAGAGAGAGAGAGAGAG 675
Qy 1056 TCTGTTGATGGAATATCTGATCAAGTTTATCTCCCTTAAACCGTATGAGAGATTTGAGA 1115
Db 676 TCTGTTGATGGAATATCTGATCAAGTTTATCTCCCTTAAACCGTATGAGAGATTTGAGA 735
Qy 1116 CCAGTCTCAAGATTAAGCTTATGAAAACTAGCTATCGGTGACACATCAATCT 1175
Db 736 CCAGTCTCAAGATTAAGCTTATGAAAACTAGCTATCGGTGACACATCAATCT 795
Qy 1176 CAAGATTTACTAGCTCAAGCAACAAAGATTTTAAACAAACCCAGGCTTATGAGTT 1235
Db 796 CAAGATTTACTAGCTCAAGCAACAAAGATTTTAAACAAACCCAGGCTTATGAGTT 855
Qy 1236 TATGAAGTGAATCTCTCAATGCTGATCATGATGATGATGATGATGATGATGATGATGAT 1295
Db 856 TATGAAGTGAATCTCTCAATGCTGATCATGATGATGATGATGATGATGATGATGATGAT 915
Qy 1296 ATGATCAAGAGTTTACTTACCGTGTAAATCGGAAACAAAGCTTATGAGATCAATTA 1355
Db 916 ATGATCAAGAGTTTACTTACCGTGTAAATCGGAAACAAAGCTTATGAGATCAATTA 975
Qy 1356 AAATCTGCTGATGAG 1415
Db 976 AAATCTGCTGATGAG 1035
Qy 1416 CTTAAAAAAGGGGAAAGCCGTATGATCCCTTATGATGATGATGATGATGATGATGATGAT 1475
Db 1036 CTTAAAAAAGGGGAAAGCCGTATGATCCCTTATGATGATGATGATGATGATGATGATGAT 1095
Qy 1476 ATCAATATCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1535
Db 1096 ATCAATATCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1155
Qy 1536 AGCGAAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1595
Db 1156 AGCGAAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1215
Qy 1596 TACAAATATCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1655
Db 1216 TACAAATATCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1275
Qy 1656 AATCAGATGACCAACCGATATCAACCGTTATATGAGGCAAGAGAGAGAGAGAGAG 1715
Db 1716 AATGCTAGCTTACATTTAGCTGTGT 1742
Qy 1742 AATGCTAGCTTACATTTAGCTGTGT 1742
Db 1336 AATGCTAGCTTACATTTAGCTGTGT 1362

RESULT 10

113215 1458 bp DNA linear PAT 26-JUL-1995
LOCUS Sequence 42 from patent US 5434073.
DEFINITION
ACCESSION 113215
VERSION 113215.1 GI:910563
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 1458)

Db 1 ATTGCTGACCTGAGTGGCTGCTAGAACCGTCCATCTGCAACACAGCCAAATTAGTTGTT 60
Qy 648 AGCGTTGCTGTACTGTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGTAAATC 707
Db 61 AGGTTGCTGTGTACTGTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGTAAAT 120
Qy 708 GATCTAACATCAAGACCTGCTCATGAGGAAAGACAGCAAGGCTTAAGTCCAAATCA 767
Db 121 GACCTAACATCAAGACCTGCTCATGAGGAAAGACAGCAAGGCTTAAGTCCAAATCA 180
Qy 768 AAACATTTGCTACTGATAGTGGGGGCAATGCACTTAACCTTGAAGAACTGACTTACTA 827
Db 181 AAACATTTGCTACTGATAGTGGGGGCAATGCACTTAACCTTGAAGAACTGACTTACTA 240
Qy 828 AAGGCTATTCAAGAACCAATTGATGCTTAACCTCAAGTAAAGCACTTCTTGAAGTC 887
Db 241 AAGGCTATTCAAGAACCAATTGATGCTTAACCTCAAGTAAAGCACTTCTTGAAGTC 300
Qy 888 ATTGATTTCGAAGCGATGCAACCATTAATGATGCAAGCGCAAGTCTTCTTGTGAC 947
Db 301 ATTGATTTCGAAGCGATGCAACCATTAATGATGCAAGCGCAAGTCTTCTTGTGAC 360
Qy 948 AAAGATGTTGCGTAACCTTGGCCGACCCCAACCTGTCGAAGATTTTGTGAAGCGACAT 1007
Db 361 AAAGATGTTGCGTAACCTTGGCCGACCCCAACCTGTCGAAGATTTTGTGAAGCGACAT 420
Qy 1008 GTGCGCGTTAGACCAATTAAGAAAGAAACCAATACAAAGCAAGGAAATCTGTGATGTG 1067
Db 421 GTGCGCGTTAGACCAATTAAGAAAGAAACCAATACAAAGCAAGGAAATCTGTGATGTG 480
Qy 1068 GAAATATCTGTACAGTTTACTCCCTTAAACCTGATGACATTTTCAAGCAGGTCTCAAA 1127
Db 481 GAAATATCTGTACAGTTTACTCCCTTAAACCTGATGACATTTTCAAGCAGGTCTCAAA 540
Qy 1128 GATATCTAGCTATTGAAAGAACTAGCTATGCTGTACACATCACTCAAGATTTACTA 1187
Db 541 GATATCTAGCTATTGAAAGAACTAGCTATGCTGTACACATCACTCAAGATTTACTA 600
Qy 1188 GCTCAAGCAAAAGCATTTTAAACAAAGAAACCAAGGCTATGAGTTTAAAGAACTGAC 1247
Db 601 GCTCAAGCAAAAGCATTTTAAACAAAGAAACCAAGGCTATGAGTTTAAAGAACTGAC 660
Qy 1248 TCCTCAATCTGCACTCATGACAAATGACATTTTCCGTAGATTTTAAAGATCAAGAG 1307
Db 661 TCCTCAATCTGCACTCATGACAAATGACATTTTCCGTAGATTTTAAAGATCAAGAG 720
Qy 1308 TTTACTTACCGGTGTTAAATTCGGGAAACAAGCTTATAGATCAATAAATTTGCTGTG 1367
Db 721 TTTACTTACCGGTGTTAAATTCGGGAAACAAGCTTATAGATCAATAAATTTGCTGTG 780
Qy 1368 AATGAAGAAATTAACCACTGACCTGATCTGTGAAATTTTACGTCTTAAAGAAAGGG 1427
Db 781 AATGAAGAAATTAACCACTGACCTGATCTGTGAAATTTTACGTCTTAAAGAAAGGG 840
Qy 1428 GAAAGACCGGTATGATCCCTTGAATCGAGTCACTTGAACCTGTCACCATGAATAGCTT 1487
Db 841 GAAAGACCGGTATGATCCCTTGAATCGAGTCACTTGAACCTGTCACCATGAATAGCTT 900
Qy 1488 GATGTGATACCAACGAATGCTTAAGAAAGTGAAGAGAGCTTAAAGAGTAAAGCTTAAC 1547
Db 901 GATGTGATACCAACGAATGCTTAAGAAAGTGAAGAGAGCTTAAAGAGTAAAGCTTAAC 960
Qy 1548 TTTGACTTCAGAGATTTTATAGATCTGTGATTAAGGCTAAAGCTTCAACAAATCTC 1607
Db 961 TTTGACTTCAGAGATTTTATAGATCTGTGATTAAGGCTAAAGCTTCAACAAATCTC 1020
Qy 1608 GATGCTTTTGTATTAAGACTATTAACCTTAACTGAAAGATAGAGATTAACAGATGAC 1667
Db 1021 GATGCTTTTGTATTAAGACTATTAACCTTAACTGAAAGATAGAGATTAACAGATGAC 1080
Qy 1668 ACCAACCGTATCATTAACCGTTTATATGAGCAAGCGAACCGGAAGAGATGCTAGCTAC 1727
Db 1081 ACCAACCGTATCATTAACCGTTTATATGAGCAAGCGAACCGGAAGAGATGCTAGCTAT 1140

Qy 1728 CATTTAGCTGTGTGT 1742
Db 1141 CATTTAGCTGTATGAT 1155

RESULT 12
LOCUS CQ797820 7057 bp DNA linear PAT 20-APR-2004
DEFINITION Sequence 26 from Patent WO2004029256.
ACCESSION CQ797820
VERSION CQ797820.1 GI:46426093
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1
AUTHORS selman-housein Sosa,G., aguiar Cabeza,E., gonz lez Quintero,A.D. and ramos gonz lez,O.
TITLE Vector for the production of transplastomic angiosperm plants
JOURNAL Patent: WO 2004029256-A 26 08-APR-2004;
CENTRO DE INGENIERIA GENETICA Y BIOTECNOLOGIA (CU)
FEATURES
source
1. 7057
Location/Qualifiers
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Description of Artificial Sequence: Nucleotide sequence of DNA fragment from the vector pVPTA-Batrep between the rice atpB and tobacco rbcL borders."

ORIGIN
Query Match 53.7%; Score 1124.6; DB 6; Length 7057;
Best Local Similarity 98.4%; Pred. No. 9.9e-292;
Matches 1136; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 588 ATTGCTGACCGTAAGTGGCTGTAGACCGTCCATGCTGCAACACAGCCAAATGTGTG 647
Db 2023 ATTGCTGACCGTAAGTGGCTGTAGACCGTCCATGCTGCAACACAGCCAAATGTGTG 2082
Qy 648 AGCGTTGCTGTACTGTGTAAGGGGACGAATCAAGACATTAGTCTTAAATTTTGTAAATC 707
Db 2083 AGCGTTGCTGTACTGTGTAAGGGGACGAATCAAGACATTAGTCTTAAATTTTGTAAAT 2142
Qy 708 GATCTAACATCAAGACCTGCTCATGAGGAAAGACAGCAAGGCTTAAGTCCAAATCA 767
Db 2143 GATCTAACATCAAGACCTGCTCATGAGGAAAGACAGCAAGGCTTAAGTCCAAATCA 2202
Qy 768 AAACCATTTGCTACTGATAGTGGCGGATGCACTTAACCTTGAAGAACTGACTTACTA 827
Db 2203 AAACCATTTGCTACTGATAGTGGCGGATGCACTTAACCTTGAAGAACTGACTTACTA 2262
Qy 828 AAGGCTATTCAAGAAACAATTGATGCTTAACGTCACAGTAAAGCACTACTTGAAGTC 887
Db 2263 AAGGCTATTCAAGAAACAATTGATGCTTAACGTCACAGTAAAGCACTACTTGAAGTC 2322
Qy 888 ATTGATTTCGAAGCGATGCAACCATTAATGATGCAAGCGCAAGTCTTCTTGTGAC 947
Db 2323 ATTGATTTCGAAGCGATGCAACCATTAATGATGCAAGCGCAAGTCTTCTTGTGAC 2382
Qy 948 AAAGATGTTGCGTAACTTGGCCGACCCCAACCTGTCGAAGATTTTGTGAAGCGACAT 1007
Db 2383 AAAGATGTTGCGTAACTTGGCCGACCCCAACCTGTCGAAGATTTTGTGAAGCGACAT 2442
Qy 1008 GTGCGCGTTAGACCAATTAAGAAAGAAACCAATACAAAGCAAGGAAATCTGTGATGTG 1067
Db 2443 GTGCGCGTTAGACCAATTAAGAAAGAAACCAATACAAAGCAAGGAAATCTGTGATGTG 2502
Qy 1068 GAATATATCTGTACAGTTTACTCCCTTAAACCTGATGACATTTTCAAGCAGGTCTCAAA 1127
Db 2503 GAATATATCTGTACAGTTTACTCCCTTAAACCTGATGACATTTTCAAGCAGGTCTCAAA 2562
Qy 1128 GATATCTAGCTATTGAAAGAACTAGCTATGCTGTGACACATCACTCAAGATTTACTA 1187

Db 2563 GATACGAGCTATGAAAAACATGAGTATCGGTGACACATCATCTCAAGAAATTACTA 2622
Qy 1188 GCTCAAGCAAAAGCATTTTAAACAAAACCCAGCGCTTACGATTTATGAAACGTCAC 1247
Db 2623 GCTCAAGCAAAAGCATTTTAAACAAAACCCAGCGCTTATGAAATTTATGAAACGTCAC 2682
Qy 1248 TCCTCATCGTCACTGATGACATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 1307
Db 2683 TCCTCATCGTCACTGATGACATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 2742
Qy 1308 TTACTTACCGGTGTTAAAAATCGGGAACAAGCTTATGATCAATAAAAATCTGCTCG 1367
Db 2743 TTACTTACCATGTCATAAATTCGGGAACAAGCTTATGATCAATAAAAATCTGCTCG 2802
Qy 1368 AATGAAGAATTAACAACACCTGACCTGATCTCTGAGAAATTTACGCTTAAAAAGGG 1427
Db 2803 AATGAAGAATTAACAACACCTGACCTGATCTCTGAGAAATTTACGCTTAAAAAGGG 2862
Qy 1428 GAAAGCCGTATGATCCCTGATCGGACCTGTAACCTGTAACCAATCAATAGCTT 1487
Db 2863 GAAAGCCGTATGATCCCTGATCGGACCTGTAACCTGTAACCAATCAATAGCTT 2922
Qy 1488 GATGCGATACCAACGATTTGTTAAAAAGTGAAGCACTCTTAACAGCTAGCGAAGTAAAC 1547
Db 2923 GATGCGATACCAACGATTTGTTAAAAAGTGAAGCAAGCTCTTAACAGCTAGCGAAGTAAAC 2982
Qy 1548 TTAGACTTCAGAGATTTATACGATCTCTGATTAAGCTTAACTTACCTTCAACATCTC 1607
Db 2983 TTAGACTTCAGAGATTTATACGATCTCTGATTAAGCTTAACTTACCTTCAACATCTC 3042
Qy 1608 GATGCTTTTGTATTTATGACCTATACCTTAACTGGAAGATGAGATTAATCAAGATGAC 1667
Db 3043 GATGCTTTTGTATTTATGACCTATACCTTAACTGGAAGATGAGATTAATCAAGATGAC 3102
Qy 1668 ACCAAGCGTATCAATACCGTTTATATGAGCAAGCAACCGAAGAGAGAAATGCTAGCTAC 1727
Db 3103 ACCAAGCGTATCAATACCGTTTATATGAGCAAGCAACCGAAGAGAGAAATGCTAGCTAT 3162
Qy 1728 CATTTAGCTGTGGGT 1742
Db 3163 CATTTAGCTGTATGAT 3177

RESULT 13
A20006 1335 bp DNA linear PAT 14-JUL-1995
LOCUS SEQ ID NO: 14; Streptokinase gene.
DEFINITION A20006
ACCESSION A20006 GI:1247579
VERSION A20006.1 GI:1247579
KEYWORDS
SOURCE
ORGANISM
Streptococcus dysgalactiae subsp. equisimilis
Streptococcus dysgalactiae subsp. equisimilis
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.
1 (bases 1 to 1335)
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1.1335
/organism="Streptococcus dysgalactiae subsp. equisimilis"
/mol_type="unassigned DNA"
/sub_species="equisimilis"
/db_xref="taxon:119602"
7.1329
/codon_start=1
/transl_table=11
/product="Streptokinase"
/protein_id="CAA01484.1"
/db_xref="GI:1247580"
/translation="MKNYLSFGMPALIFALTPTGVNSVOAIAQPEWLLDPSVNSQL
VNSVAGTVBGTNDISLKPFEIDLTSPAHGKTBOGLSPKSPFATDSCAMPKLEK

ADILKAIQEOILIANVHNSNDYFEVYIDPASPATITDRNGKYVFAUDKSGSVTLPTQPVOR
FLISGVRVRYKEKPIONOKASVDPVETVOFTEPLNDPDDPRLKDTKLTAIGD
TITSOELAAQASITLNTHTGTYTVEEDSIVTHDNDIPEITLPMDOEFYVNRBO
AYEINKSGABEINNTDILSEKTYVKKGEKPPDPDRSLKLFITYDVYVNBELL
KSEQLTASERNLDFRDIYDRDAKLLYNLDAFIMDYLTJTKVEVDNHDNRLIT
VYMKRDEGENASYHLYVDORYTEBEREVSYLRYGTPIPDNPNDK"

ORIGIN

Query Match 53.5%; Score 1122.4; DB 6; Length 1335;
Best Local Similarity 98.2%; Pred. No. 3.4e-291;
Matches 1135; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
Qy 587 TATTGCTGACCTAGTGGCTGTAGACCGGCACTGTCTCAACAACAGCCAAATGGTGT 646
Db 84 TATTGCTGACCTAGTGGCTGTAGACCGGCACTGTCTCAACAACAGCCAAATGGTGT 143
Qy 647 TAGCGTGTGTTACTGTTGAGGGGACGAATCAAGAATTAGCTTAAATTTTGAAT 706
Db 144 TAGCGTGTGTTACTGTTGAGGGGACGAATCAAGAATTAGCTTAAATTTTGAAT 203
Qy 707 CGATCTAACATCAAGACTGCTCATGAGAGAAAGCAGAGAGCTTAAGTCCAAATC 766
Db 204 TGACCTAACATCAAGACTGCTCATGAGAGAAAGCAGAGAGCTTAAGTCCAAATC 263
Qy 767 AAAACCAATTCCTATGATAGTGGCGGATGTCAATAACTTGAGAAAGCTGACTACT 826
Db 264 AAAACCAATTCCTATGATAGTGGCGGATGTCAATAACTTGAGAAAGCTGACTACT 323
Qy 827 AAAAGCTATTCAGAAACAATGATGCTTAACGTCACAGTAACGACCTTATGAGGT 886
Db 324 AAAAGCTATTCAGAAACAATGATGCTTAACGTCACAGTAACGACCTTATGAGGT 383
Qy 887 CATTGATTTTGCAAGCGATGCAACCAATTACTGATGCAAGCGCAAGGCTTACTTGTCTGA 946
Db 384 CATTGATTTTGCAAGCGATGCAACCAATTACTGATGCAAGCGCAAGGCTTACTTGTCTGA 443
Qy 947 CAAGATGTTGGTGAACCTTGGCCGACCAACCTGTCCAAAGATTTTGTCTAGCGGACA 1006
Db 444 CAAGATGTTGGTGAACCTTGGCCGACCAACCTGTCCAAAGATTTTGTCTAGCGGACA 503
Qy 1007 TGTGCGGTTTGAACCATATTAAGAAAAACAATCAAAAACCAAGCGAATCTGTGATGT 1066
Db 504 TGTGCGGTTTGAACCATATTAAGAAAAACAATCAAAAACCAAGCGAATCTGTGATGT 563
Qy 1067 GGAATATATCTGTACATTTACTCCCTTAACTCTGATGACATTTCAAGACAGTCTGAA 1126
Db 564 GGAATATATCTGTACATTTACTCCCTTAACTCTGATGACATTTCAAGACAGTCTGAA 623
Qy 1127 AGATATCAAGCTATTTGAAAAACATAGCTATGCGTGAACAATCAATCTCAAGAAATTA 1186
Db 624 AGATATCAAGCTATTTGAAAAACATAGCTATGCGTGAACAATCAATCTCAAGAAATTA 683
Qy 1187 AGCTCAAGCAAAAGCATTTTAAACAAAACCAACCCAGGCTATGATTTGAACTGGA 1246
Db 684 AGCTCAAGCAAAAGCATTTTAAACAAAACCAACCCAGGCTATGATTTGAACTGGA 743
Qy 1247 CTCTCAATGCTCACTATGACAAATGACATTTTCGTATGCAATTTTCAATGATGACGA 1306
Db 744 CTCTCAATGCTCACTATGACAAATGACATTTTCGTATGCAATTTTCAATGATGACGA 803
Qy 1307 GTTTACTTACCGGTGTTAAAAATTCGGGAACAAGCTTATAGATCAATAAAAATCTGCT 1366
Db 804 GTTTACTTACCGGTGTTAAAAATTCGGGAACAAGCTTATAGATCAATAAAAATCTGCT 863
Qy 1367 GATGAAGAATTAACAACACCTGACCTGATCTGAGAAATTTACGCTTTAAAAAGG 1426
Db 864 GATGAAGAATTAACAACACCTGACCTGATCTGAGAAATTTACGCTTTAAAAAGG 923
Qy 1427 GAAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCACATCAATAGCT 1486
Db 924 GAAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCACATCAATAGCT 983
Qy 1487 TGATGTGATTCACACGAATTCCTTAAAAAGTGAAGACGCTTTAACAGCTAGCGAAGTAA 1546

Db	964	TGATGTCAACACCAACGAATTCCTAAAAAGCAGACGCTTTAAACGCTAGGAAACGTAA	1043
QY	1547	CTTAACTTCAGAGATTATATACGATCTCTGTATTAAGCTAAACTCTAACAACTCT	1606
Db	1044	CTTAACTTCAGAGATTATATACGATCTCTGTATTAAGCTAAACTCTAACAACTCT	1103
QY	1607	CGATGCTTTGGTATTATAGACTATTAACCTTAACTGGAAGAAAGTAGAGTAATCAAGTGA	1666
Db	1104	CGATGCTTTGGTATTATAGACTATTAACCTTAACTGGAAGAAAGTAGAGTAATCAAGTGA	1163
QY	1667	CACCAACCGTATCATTAACCGTTATATATGAGCAAGCAACCGAAGGAGAAATGTAGCTA	1726
Db	1164	CACCAACCGTATCATTAACCGTTATATATGAGCAAGCAACCGAAGGAGAAATGTAGCTA	1223
QY	1727	CCATTAGCTGGTGT 1742	
Db	1224	TCATTAGCTGATGAT 1239	
RESULT 14			
LOCUS	113194	1335 bp	linear
DEFINITION	Sequence 14 from patent US 5434073.		
ACCESSION	113194		
VERSION	113194.1	GI:910542	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 1335)		
AUTHORS	Dawson, K., Hunter, M. G. and Czaplinski, L. G.		
TITLE	Fluorimetric and anti-thrombotic cleavable dimers		
JOURNAL	Patent: US 5434073-A 14 18-JUL-1995;		
FEATURES	Location/Qualifiers		
source	1..1335		
	/organism="unknown"		
	/mol_type="unassigned DNA"		
ORIGIN			
Query Match	53.5%; Score 1122.4; DB 6; Length 1335;		
Best Local Similarity	98.2%; Pred. No. 3.4e-291;		
Matches 1135; Conservative	0; Mismatches 21; Indels 0; Gaps 0;		
QY	587	TATTGCTGACCTGAGTGAGTGCTGATGACCGTCCATCTGTCAACAAACGCAATTTGTTGT	646
Db	84	TATTGCTGACCTGAGTGAGTGCTGATGACCGTCCATCTGTCAACAAACGCAATTTGTTGT	143
QY	647	TAGGCTGCTGCTACTGTTGAGAGGAGCAACATCAACACTTGTCTTAAATTTTGTGAAT	706
Db	144	TAGGCTGCTGCTACTGTTGAGAGGAGCAACATCAACACTTGTCTTAAATTTTGTGAAT	203
QY	707	CGATTTACATCAACGACCTGCTCATGTAGAGAAAGCAGAGCAAGGCTTAAGTCCAAATC	766
Db	204	TGACTTAACATCAACGACCTGCTCATGTAGAGAAAGCAGAGCAAGGCTTAAGTCCAAATC	263
QY	767	AAAACGATTTGCTACTGATAGTGGCGCATGTCACTTAACTTGAAAGCTGACTTACT	826
Db	264	AAAACGATTTGCTACTGATAGTGGCGCATGTCACTTAACTTGAAAGCTGACTTACT	323
QY	827	AAAGGCTATTCAAGAACAATTGATGCTGAAGTCCACGTTAACGACGACTTTGAGGT	886
Db	324	AAAGGCTATTCAAGAACAATTGATGCTGAAGTCCACGTTAACGACGACTTTGAGGT	383
QY	887	CATTGATTTTGAAGCGATGACCACTAATGATGAGAAAGGCAAGGTCTACTTTGCTGA	946
Db	384	CATTGATTTTGAAGCGATGACCACTAATGATGAGAAAGGCAAGGTCTACTTTGCTGA	443
QY	947	CAAGATGTTTCGTTAACCCTTGGCCGACCAACTGTCCAGAAATTTTGTCTAACGAGCA	1006
Db	444	CAAGATGTTTCGTTAACCCTTGGCCGACCAACTGTCCAGAAATTTTGTCTAACGAGCA	503
QY	1007	TGTGGCGCTTGAACCATATTAAGAAAAACAATTAACAAAAACGAGAAATCTGTTGATG	1066

[illegible]

```

/codon_start=1
/translate=11
/product="streptokinase fused to yeast alpha-factor"
/protein_id="CA01487.1"
/db_xref="GI:1247851"
/translation="WRPSP1FTAVLPAASSALAPVNTTEDETAQIPAEAVIGYLDL
EGPDFVAVLPFSNNTNGNLFINTTASIAKEBGSVSLDKRIAPEMILDPSPVNSQ
LVSVAGTEGTMQDLSLKFEIDLTSPAPAGKTEQGLSPKSPAPATDSGAMPKLE
KADLKAIQOLIANVSNNDYEPVIDPASPAITTDNGKYPAPDKGSAVLPPOPIO
EPILSGHVRVRYKPKRIQONAKSVDPYTPQTPPLNDDPFRGLDKTLKTLAIG
DITTSQELAQOSILNKTHPGYTYERDSSIYHNDIFPTLI.PMQEFTYHAKNE
QAYEINKSGSLNEINNTDLSKEYVYALKKEKPYDPPDRSHLFLPIKYDVNTNEL
LKESQULTASERINDFRDLDPDRKALLYNLDAPFGMDYTLGKXEDNDDPTNRIL
TYVMGKRPBENASVHLAYDKRITREEREVYSYLRTGTPIPNPNDK"

```

ORIGIN

```

Query Match      53.5%; Score 1122.4; DB 6; Length 1512;
Best Local Similarity 97.8%; Pred. No. 3.4e-291;
Matches 1138; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 579 GATGTTGATATGCTGAGACCTGAGTGGCTGCTAGACCTCATCTGTCAACAACAGCCAA 638
DB 253 GATTAAGAAATGCTGAGACCTGAGTGGCTGCTAGACCTCATCTGTCAACAACAGCCAA 312
QY 639 TTGGTTGTAGCGTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 698
DB 313 TTAGTTGTAGCGTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 372
QY 699 TTTGAAATCGATCTAATCAACGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 758
DB 373 TTTGAAATCGATCTAATCAACGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 432
QY 759 CCAAAATCAAAACCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 818
DB 433 CCAAAATCAAAACCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 492
QY 819 GACTTACTAAAGGCTATTCAAGAACATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 878
DB 493 GACTTACTAAAGGCTATTCAAGAACATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 552
QY 879 TTTGAGGTCAATTTGTCAGAGGATGCAACCATTAATGATCGAAGCGCAAGGTCTAC 938
DB 553 TTTGAGGTCAATTTGTCAGAGGATGCAACCATTAATGATCGAAGCGCAAGGTCTAC 612
QY 939 TTTGCTGCAAAAGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 998
DB 613 TTTGCTGCAAAAGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 672
QY 999 AGCGGACATGCGCGCTTACGACATATTAAGAAACCAATACAAACCAAGCGAAATCT 1058
DB 673 AGCGGACATGCGCGCTTACGACATATTAAGAAACCAATACAAACCAAGCGAAATCT 732
QY 1059 GTTGATGTGAATTTACTGTATCAAGTTTACTCTTTAAACCTTGATGACGATTTGACACCA 1118
DB 733 GTTGATGTGAATTTACTGTATCAAGTTTACTCTTTAAACCTTGATGACGATTTGACACCA 792
QY 1119 GGTCTCAAAGATTAAGTATTTGAAAACTAGCTATCGGTGACACCATCAATCTCAA 1178
DB 793 GGTCTCAAAGATTAAGTATTTGAAAACTAGCTATCGGTGACACCATCAATCTCAA 852
QY 1179 GAATTAAGTCTAAGCAAGCAAGATTTTAAACAAACCAAGGCTATAGCAATTTAT 1238
DB 853 GAATTAAGTCTAAGCAAGCAAGATTTTAAACAAACCAAGGCTATAGCAATTTAT 912
QY 1239 GAAAGTACTCTCAATCTGATCATGATGACAAATTTTCCGTACGATTTTACCAATG 1298
DB 913 GAAAGTACTCTCAATCTGATCATGATGACAAATTTTCCGTACGATTTTACCAATG 972
QY 1299 GATCAAGATTTTACTTACCGGTAAAAAATCGGAAACAAAGCTTATAGGATCAATAAAAA 1358
DB 973 GATCAAGATTTTACTTACCGGTAAAAAATCGGAAACAAAGCTTATAGGATCAATAAAAA 1032
QY 1359 TCTGCTGTGAATGAAGAAATTAACAACGACCTGATCTGTGAAAAATATTAGTCTT 1418

```

```

DB 1033 TCTGCTGTGAATGAAGAAATTAACAACGACCTGATCTGTGAAAAATATTAGTCTT 1092
QY 1419 AAAAAAGGGAAGAAAGCCGTATGATCCCTTTATGTGAGTCACTTGAACCTGTACCATC 1478
DB 1093 AAAAAAGGGAAGAAAGCCGTATGATCCCTTTATGTGAGTCACTTGAACCTGTACCATC 1152
QY 1479 AAATACGTGATGTGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAACAGCTAGC 1538
DB 1153 AAATACGTGATGTGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAACAGCTAGC 1212
QY 1539 GAACTAACTTGAACCTTCAAGATTTATACGATCTCGTGAATTAAGGCTTAACTACTTAC 1598
DB 1213 GAACTAACTTGAACCTTCAAGATTTATACGATCTCGTGAATTAAGGCTTAACTACTTAC 1272
QY 1599 AACATCTCGATGCTTTTGTATTTATGACTATACCTTAACCTGAAAAATAGAGATAT 1658
DB 1273 AACATCTCGATGCTTTTGTATTTATGACTATACCTTAACCTGAAAAATAGAGATAT 1332
QY 1659 CACGATGACACCAACCGATATACGCTTTATATGAGGCAAGCGAACGAGAGAGAT 1718
DB 1333 CACGATGACACCAACCGATATACGCTTTATATGAGGCAAGCGAACGAGAGAGAT 1392
QY 1719 GCTAGCTAACATTTAGCTGTGCT 1742
DB 1393 GCTAGCTAACATTTAGCTGTGCT 1416

```

Search completed: February 1, 2006, 04:17:49
 Job time : 10656.6 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2006, 10:44:08 ; Search time 1267.07 Seconds
(without alignments)
11024.772 Million cell updates/sec

Title: US-09-940-235-12

Perfect score: 2096
Sequence: 1 cgaagaccatcatcgtgtc.....cctcaccgcatcgtcttag 2096

Scoring table: IDENTITY NUC
Gapop 10'-0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 10%
Listing first 45 summaries

Database : N_Geneseq_21:*

- 1: geneseqn1980s:*
- 2: geneseqn1990s:*
- 3: geneseqn2000s:*
- 4: geneseqn2001as:*
- 5: geneseqn2001bs:*
- 6: geneseqn2002as:*
- 7: geneseqn2002bs:*
- 8: geneseqn2003as:*
- 9: geneseqn2003bs:*
- 10: geneseqn2003cs:*
- 11: geneseqn2003ds:*
- 12: geneseqn2004as:*
- 13: geneseqn2004bs:*
- 14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2096	100.0	2096	3	AAA37643
2	1684	80.3	1782	3	AAA37642
3	1496.2	71.4	1541	3	AAA37644
4	1165.8	55.6	1661	3	AAA37637
5	1150.8	54.9	2385	2	AAx80497
6	1147	54.7	1377	3	AAA37622
7	1145.4	54.6	1242	2	AAx80492
8	1145.4	54.6	1245	3	AAA37633
9	1145.4	54.6	1254	3	ABA05546
10	1145.4	54.6	8893	6	ABA05547
11	1143.8	54.6	1242	2	AAx16632
12	1143.8	54.6	2030	2	AAQ11651
13	1140.6	54.1	1242	2	AAx16633
14	1134.2	54.1	1327	3	AAA37628
15	1132.4	54.0	2366	2	AAx77778
16	1125.4	53.7	1458	2	AAQ12162
17	1124.6	53.7	7057	12	ADMO1294
18	1122.4	53.5	1335	2	AAQ12156
19	1122.4	53.5	1512	2	AAQ12158

20	1122.2	53.5	2589	2	AAQ12160	AAQ12160	OMPAL str
21	1121.4	53.5	1245	2	AAQ20665	AAQ20665	SKC-2 str
22	1121.4	53.5	1467	2	AAQ12490	AAQ12490	Factor Xa
23	1116.6	53.3	1407	1	AAx70106	AAx70106	DNA encod
24	1114.4	53.2	1473	2	AAQ05603	AAQ05603	Streptoki
25	1106.4	52.8	2568	1	AAx50493	AAx50493	Sequence
26	1104.8	52.7	1323	2	AAx29961	AAx29961	Vector pS
27	1096.2	52.3	1242	5	AAx82144	AAx82144	Mutant st
28	1093.6	52.2	1122	3	AAx299251	AAx299251	DNA encod
29	1093.6	52.2	1158	3	AAx299252	AAx299252	DNA encod
30	1086	51.8	1209	3	AAx299249	AAx299249	DNA encod
31	1086	51.8	1245	3	AAx299250	AAx299250	DNA encod
32	1077.2	51.4	2253	2	AAQ12161	AAQ12161	Met-core
33	1076	51.3	1119	2	AAQ12159	AAQ12159	Truncated
34	996	47.5	1320	6	ABx70192	ABx70192	Streptoco
35	996	47.5	1323	13	ADx83811	ADx83811	S. pyogen
36	991.2	47.3	1473	2	AAQ05604	AAQ05604	Streptoki
37	973	46.4	1068	2	AAx80493	AAx80493	Recombina
38	968.4	46.2	2208	2	AAx83589	AAx83589	Recombina
39	899.4	42.9	1245	10	ADx48644	ADx48644	Streptoco
40	759.4	36.2	1262	2	AAQ10230	AAQ10230	Synthetic
41	450	21.5	450	2	AAx80494	AAx80494	Recombina
42	330	15.7	777	3	AAx37632	AAx37632	Human fib
43	330	15.7	1929	13	ADx17488	ADx17488	Nucleotid
44	330	15.7	1929	13	ADx97657	ADx97657	Human fib
45	330	15.7	1929	14	ADx44479	ADx44479	Human fib

ALIGNMENTS

RESULT 1	AAA37643	standard; DNA; 2096 BP.
ID	AAA37643	
XX	AAA37643;	
AC	15-SEP-2003 (revised)	
DT	13-OCT-2000 (first entry)	
XX	Chimeric SK-FBD coding sequence.	
DE	Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;	
KW	plasminogen; human; fibrinectin; thrombolytic therapy;	
KW	cardiovascular disorder; ss.	
XX	Streptococcus dysgalactiae subsp. equisimilis.	
OS	Homo sapiens.	
OS	Chimeric.	
PN	EP1024192-A2.	
XX	02-AUG-2000.	
PD	23-DEC-1999; 99BP-00310541.	
PF	24-DEC-1998; 98IN-DE003825.	
XX	(COUL) CSIR COUNCIL SCI IND RES.	
PA	Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;	
PI	Yadav M;	
PI	WPI, 2000-516032/47.	
DR	Hybrid streptokinase-fibrin binding domain polypeptides useful for	
XX	thrombolytic therapy comprises a streptokinase fused with fibrin binding	
PT	domains of human fibronectin.	
PT	Example 6; Fig 22b; 58bp; English.	
PS	This sequence represents a chimeric streptokinase-fibrin binding domain	
XX	(SK-FBD) protein coding sequence. The invention relates to a hybrid	
CC		

CC plasminogen activator (PA) comprises a polypeptide fusion between
CC streptokinase (SK), which are capable of plasminogen (PG) activation, and
CC fibrin binding regions of human fibronectin, which are from fibrin
CC binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA possesses the
CC ability to bind with fibrin independently and also characteristically
CC retains a PG activation ability which becomes evident only after a
CC pronounced duration, or lag, after exposure of the PA to a suitable
CC animal or human PG. The hybrid streptokinase-fibrin binding domain
CC polypeptides are useful in thrombolytic therapy for various kinds of
CC cardiovascular disorders. The hybrids have enhanced fibrin selectivity as
CC well as kinetics of plasminogen activation that are distinct from that of
CC natural streptokinase in being characterised by a temporary delay, or lag
CC of several minutes in the natural rate of the catalytic conversion of
CC plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins
CC can bind tightly with fibrin in blood clots soon after introduction into
CC the vascular system without significantly activating the circulating
CC plasminogen to plasmin, thus aiding in the localisation of the
CC plasminogen activation process to the site of pathological thrombus. This
CC overcomes systemic plasminogen activation encountered during clinical use
CC of streptokinase. (Updated on 15-SEP-2003 to standardise OS field)

XX Sequence 2096 BP; 643 A; 480 C; 484 G; 489 T; 0 U; 0 Other;

Query Match 100.0%; Score 2096; DB 3; Length 2096;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2096; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGAAGACCATTCATGTTGTTGCTCGAGTCGAGCGTTTTCAGACAGCTGCTTCAGC 60
DB 1 CGAAGACCATTCATGTTGTTGCTCGAGTCGAGCGTTTTCAGACAGCTGCTTCAGC 60
QY 61 TTGCGTCGCGTATCGGATGTTCAATTCGCTAACAGTAAGGCAACCCCGCAGCTTAGCC 120
DB 61 TTGCGTCGCGTATCGGATGTTCAATTCGCTAACAGTAAGGCAACCCCGCAGCTTAGCC 120
QY 121 GGGTCTCAACGACAGGACGATCATGCGCACCCGTGGCAGAGCCCAACGCTGCGCG 180
DB 121 GGGTCTCAACGACAGGACGATCATGCGCACCCGTGGCAGAGCCCAACGCTGCGCG 180
QY 181 AGATCTCGATCCGCGAATTAATACGACTCATATAGGAGAACCAACGGTTTCCCTC 240
DB 181 AGATCTCGATCCGCGAATTAATACGACTCATATAGGAGAACCAACGGTTTCCCTC 240
QY 241 TAGAATAATTTTGTTTAACTTTAAGAGAGATATCCATGCTGCAAGCAACAATTT 300
DB 241 TAGAATAATTTTGTTTAACTTTAAGAGAGATATCCATGCTGCAAGCAACAATTT 300
QY 301 GTACCATAGCTGAGAAAGTGTTTTATCATGCTGCTGGGAACTTCTATGCTGAGGAA 360
DB 301 GTACCATAGCTGAGAAAGTGTTTTATCATGCTGCTGGGAACTTCTATGCTGAGGAA 360
QY 361 ACGTGGAGAGAGGACGAGGACGATCACTTGCACTTTAGAAATAGATGACAGATCAG 420
DB 361 ACGTGGAGAGAGGACGAGGACGATCACTTGCACTTTAGAAATAGATGACAGATCAG 420
QY 421 ACACAGAGACATCTTATGATTTGAGACACCTGAGCAAGAGAAATATGAGGAAACC 480
DB 421 ACACAGAGACATCTTATGATTTGAGACACCTGAGCAAGAGAAATATGAGGAAACC 480
QY 481 TGCTCAGAGCATCTGCAAGGCAAGGCGAGAGAGTGAATGTGAGAGGACACCT 540
DB 481 TGCTCAGAGCATCTGCAAGGCAAGGCGAGAGAGTGAATGTGAGAGGACACCT 540
QY 541 CTGTGAGACACATCAGAGCGAGCTTGCCCTTCAACCGATGTTCTGATTTCTGACCTG 600
DB 541 CTGTGAGACACATCAGAGCGAGCTTGCCCTTCAACCGATGTTCTGATTTCTGACCTG 600
QY 601 AGTGGCTGCTAGACCGTTCATCTGTCAACACAGCCAAATGTTGTTAGGCTGCTGTA 660
DB 601 AGTGGCTGCTAGACCGTTCATCTGTCAACACAGCCAAATGTTGTTAGGCTGCTGTA 660
QY 661 CTGTTAGGGGAGAAATCAAGACATTAAGCTTAAATTTTGAATGATCTAACATCAC 720
DB 661 CTGTTAGGGGAGAAATCAAGACATTAAGCTTAAATTTTGAATGATCTAACATCAC 720

DB 661 CTGTTAGGGGAGAAATCAAGACATTAAGCTTAAATTTTGAATGATCTAACATCAC 720
QY 721 GACCTGCTCAGAGGAGAAAGACAGAGCAAGGCTTAAAGCCAAATCAAAATCAATTTGCTA 780
DB 721 GACCTGCTCAGAGGAGAAAGACAGAGCAAGGCTTAAAGCCAAATCAAAATCAATTTGCTA 780
QY 781 CTGATAGGGGCGCATGTCATATAACTTGAGAAAGCTGACTTAAAGGCTATTTCAG 840
DB 781 CTGATAGGGGCGCATGTCATATAACTTGAGAAAGCTGACTTAAAGGCTATTTCAG 840
QY 841 AACAAATGATGCTAACGTCACAGTAACGACGACTACTTGAAGGTCAATTTTGCA 900
DB 841 AACAAATGATGCTAACGTCACAGTAACGACGACTACTTGAAGGTCAATTTTGCA 900
QY 841 AACAAATGATGCTAACGTCACAGTAACGACGACTACTTGAAGGTCAATTTTGCA 900
DB 841 AACAAATGATGCTAACGTCACAGTAACGACGACTACTTGAAGGTCAATTTTGCA 900
QY 901 GCGATGCAACATTAATGATGAAAGCGCAAGGTCTACTTGTCTGACAAAGATGCTGG 960
DB 901 GCGATGCAACATTAATGATGAAAGCGCAAGGTCTACTTGTCTGACAAAGATGCTGG 960
QY 961 TTAACCTTGCAGACCAACCTGTCCAGAAATTTTGTGAGGAGCATGTGGCGCTTAGC 1020
DB 961 TTAACCTTGCAGACCAACCTGTCCAGAAATTTTGTGAGGAGCATGTGGCGCTTAGC 1020
QY 1021 CATATPAAAGAAACCAATACAAAACCAAGGAAATCTGTGATGTAATATACCTGAC 1080
DB 1021 CATATPAAAGAAACCAATACAAAACCAAGGAAATCTGTGATGTAATATACCTGAC 1080
QY 1081 AGTTTACTCCCTTAAACCTGATGACGATTTTCAACGAGCTCAAAATATTAAGCTAT 1140
DB 1081 AGTTTACTCCCTTAAACCTGATGACGATTTTCAACGAGCTCAAAATATTAAGCTAT 1140
QY 1141 TGAAGAACCTAGCTATGAGTGAACCAATCATCTCAAGATTTACTCAAGACAAA 1200
DB 1141 TGAAGAACCTAGCTATGAGTGAACCAATCATCTCAAGATTTACTCAAGACAAA 1200
QY 1201 GCATTTTAAACAAACCAACCAAGGCTTAAAGATTTATGAAAGTCACTCTCAATGCTCA 1260
DB 1201 GCATTTTAAACAAACCAACCAAGGCTTAAAGATTTATGAAAGTCACTCTCAATGCTCA 1260
QY 1261 CTGATGACATGACATTTTCCGTACGATTTTACCAATGATCAAGAGTTTACTTCCGTG 1320
DB 1261 CTGATGACATGACATTTTCCGTACGATTTTACCAATGATCAAGAGTTTACTTCCGTG 1320
QY 1321 TTTAAATTCGGGAAACAGCTTATAGATGCAATTTAAATCTGTGTAAGAAATTA 1380
DB 1321 TTTAAATTCGGGAAACAGCTTATAGATGCAATTTAAATCTGTGTAAGAAATTA 1380
QY 1381 ACAACACTGACCTGATCTCTGAGAAATTTAGCTCTTAAATTTAAAGGAGGAAAGCCGTATG 1440
DB 1381 ACAACACTGACCTGATCTCTGAGAAATTTAGCTCTTAAATTTAAAGGAGGAAAGCCGTATG 1440
QY 1441 ATCCCTTTGATGCAAGTCACTTGAACCTGTTCAACCATCAAAATACGTTGATGATACCA 1500
DB 1441 ATCCCTTTGATGCAAGTCACTTGAACCTGTTCAACCATCAAAATACGTTGATGATACCA 1500
QY 1501 ACGAATGCTTAAAGATGAGAGGCTTTTAAAGCTTAAAGGAGGAAAGCTTAAAGCTTAAAG 1560
DB 1501 ACGAATGCTTAAAGATGAGAGGCTTTTAAAGCTTAAAGGAGGAAAGCTTAAAGCTTAAAG 1560
QY 1561 ATTATATGATGCTCTGATGTAAGGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAG 1620
DB 1561 ATTATATGATGCTCTGATGTAAGGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAG 1620
QY 1621 TTATGACTATACCTTAACTGAGAAAGTGAAGATTAACAGATGACACCAACCGTATAC 1680
DB 1621 TTATGACTATACCTTAACTGAGAAAGTGAAGATTAACAGATGACACCAACCGTATAC 1680
QY 1681 TTAACCTTTATATGAGGCAAGGACCCGAGAGAGATGCTAAGCTTAAAGCTTAAAG 1740
DB 1681 TTAACCTTTATATGAGGCAAGGACCCGAGAGAGATGCTAAGCTTAAAGCTTAAAG 1740
QY 1741 GTGGCAGGGGCAACAGATTGATCCATAGGCTGGAAGTGTGTTGATCATGCTGCTGGGA 1800
DB 1741 GTGGCAGGGGCAACAGATTGATCCATAGGCTGGAAGTGTGTTGATCATGCTGCTGGGA 1800


```

QY 891 GATTTGCAAGCATGCAACCATTAATGATGAAACGGCAAGGTCTACTTTGCTGACAA 950
    |||||
DB 841 GATTTGCAAGCATGCAACCATTAATGATGAAACGGCAAGGTCTACTTTGCTGACAA 900
QY 951 GATGTTGGTAACTTTGCGGACCACTGTCAGAAATTTTGTAAAGCGGACATGTG 1010
    |||||
DB 901 GATGTTGGTAACTTTGCGGACCACTGTCAGAAATTTTGTAAAGCGGACATGTG 960
QY 1011 CGGTTAGACATTAATAAGAAAAACAATCAAAACCAAGGAAATCTGTTGATGGA 1070
    |||||
DB 961 CGGTTAGACATTAATAAGAAAAACAATCAAAACCAAGGAAATCTGTTGATGGA 1020
QY 1071 TAACTGTACAGTTACTCCCTTAAACCTGATGACATTTTCAACACAGGCTCAAAAGAT 1130
    |||||
DB 1021 TAACTGTACAGTTACTCCCTTAAACCTGATGACATTTTCAACACAGGCTCAAAAGAT 1080
QY 1131 ACTAAGCTATTTGAAACCACTAGCTATGATGACACATCATCTCAAGATTAATAGCT 1190
    |||||
DB 1081 ACTAAGCTATTTGAAACCACTAGCTATGATGACACATCATCTCAAGATTAATAGCT 1140
QY 1191 CAAGCAAAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTATGAAGTCTCC 1250
    |||||
DB 1141 CAAGCAAAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTATGAAGTCTCC 1200
QY 1251 TCAATGTCATGATGACATGACATTTTCCGTAACATTTTCAAAAGGATGATGAT 1310
    |||||
DB 1201 TCAATGTCATGATGACATGACATTTTCCGTAACATTTTCAAAAGGATGATGAT 1260
QY 1311 ACTTACCGTGTAAATAATCGGGAACAAGCTTATAGATCAATATAAAATCTGCTGAT 1370
    |||||
DB 1261 ACTTACCGTGTAAATAATCGGGAACAAGCTTATAGATCAATATAAAATCTGCTGAT 1320
QY 1371 GAAGAAATATAACAACACTGACCTGATCTGGAAGATTTTACGCTTTAAAAAGGGAA 1430
    |||||
DB 1321 GAAGAAATATAACAACACTGACCTGATCTGGAAGATTTTACGCTTTAAAAAGGGAA 1380
QY 1431 AAGCGGTATGATCCCTTTGATCGGACGACCTTGAAACCTGTTCAACATCAATATGTTGAT 1490
    |||||
DB 1381 AAGCGGTATGATCCCTTTGATCGGACGACCTTGAAACCTGTTCAACATCAATATGTTGAT 1440
QY 1491 GTGATATCAACAAGATGCTTAAAGAGAGAGAGCTTTAAACAGCTAGGAAAGCTTA 1550
    |||||
DB 1441 GTGATATCAACAAGATGCTTAAAGAGAGAGAGCTTTAAACAGCTAGGAAAGCTTA 1500
QY 1551 GACTTCAGAGATTTATACGATCTCGTATAGGCTTAACTACTCTCAACAATCTCGAT 1610
    |||||
DB 1501 GACTTCAGAGATTTATACGATCTCGTATAGGCTTAACTACTCTCAACAATCTCGAT 1560
QY 1611 GCTTTGGTATTTATGACCTTAATCTTAACTGAAAGAGATTAATCAGATGACACC 1670
    |||||
DB 1561 GCTTTGGTATTTATGACCTTAATCTTAACTGAAAGAGATTAATCAGATGACACC 1620
QY 1671 AACCGTATCAAAACCGTTTATATAGGGAAGGACCCGAAGGAGAAATGCTAGCTACAT 1730
    |||||
DB 1621 AACCGTATCAAAACCGTTTATATAGGGAAGGACCCGAAGGAGAAATGCTAGCTACAT 1680
QY 1731 TTAGCTGTGTGT 1742
    |||||
DB 1681 TTAGCTGTGTGT 1692

```

RESULT 3
AAA37644
ID AAA37644 standard; DNA; 1541 BP.
XX

AC AAA37644;

DT 15-SEP-2003 (revised)
DT 13-OCT-2000 (first entry)

DE Chimeric SK-FBD coding sequence.

XX

```

KW Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
KW plasminogen; human; fibrinectin; thrombolytic therapy;
KW cardiovascular disorder; ss.
OS Streptococcus dysgalactiae subsp. equisimilis.
OS Homo sapiens.
OS Chimeric.
PN EPI024192-A2.
PD 02-AUG-2000.
PF 23-DEC-1999; 99EP-00310541.
PR 24-DEC-1998; 98IN-DE003825.
XX (COUL ) CSIR COUNCIL SCI IND RES.
PA
PI Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
PI Yadav M;
DR WPI; 2000-516032/47.
XX
XX Hybrid streptokinase-fibrin binding domain polypeptides useful for
PT thrombolytic therapy comprises a streptokinase fused with fibrin binding
PT domains of human fibrinectin.
PS Disclosure; Fig 17b; 58pp; English.
XX
XX This sequence represents a chimeric streptokinase-fibrin binding domain
CC (SK-FBD) protein coding sequence. The invention relates to a hybrid
CC plasminogen activator (PA) comprises a polypeptide fusion between
CC streptokinase (SK), which are capable of plasminogen (PG) activation, and
CC fibrin binding regions of human fibrinectin, which are from fibrin
CC binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA possesses the
CC ability to bind with fibrin independently and also characteristically
CC retains a PG activation ability which becomes evident only after a
CC pronounced duration, or lag, after exposure of the PA to a suitable
CC animal or human PG. The hybrid streptokinase-fibrin binding domain
CC polypeptides are useful in thrombolytic therapy for various kinds of
CC cardiovascular disorders. The hybrid have enhanced fibrin selectivity as
CC well as kinetics of plasminogen activation that are distinct from that of
CC natural streptokinase in being characterised by a temporary delay, or lag
CC of several minutes in the natural rate of the catalytic conversion of
CC plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins
CC can bind tightly with fibrin in blood clots soon after introduction into
CC the vascular system without significantly activating the circulating
CC blood plasminogen to plasmin, thus aiding in the localisation of the
CC plasminogen activation process to the site of pathological thrombus. This
CC overcomes systemic plasminogen activation encountered during clinical use
CC of streptokinase. (Updated on 15-SEP-2003 to standardise OS field)
CC
XX
SQ Sequence 1541 BP; 497 A; 328 C; 335 G; 381 T; 0 U; 0 Other;
Query Match 71.4%; Score 1496.2; DB 3; Length 1541;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1501; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 588 ATTGCTGACCTGAGTGGCTGCTAGACCGTCAATCTGTCAACAACAGCAATTTGTTGT 647
    |||||
DB 33 ATAGCTGTCTCTGAATGGCTACTAGATGTCTCTTCTGTAAATTAACAGCAATTTGTTGT 92
QY 648 AGCGTTGCTGTAATCTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 707
    |||||
DB 93 AGCGTTGCTGTAATCTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 152
QY 708 GATCTAATCAACGACCTGCTCATGAGGAAAGAGCAAGGCTTAAGTCCAAATTA 767
    |||||
DB 153 GATCTAATCAACGACCTGCTCATGAGGAAAGAGCAAGGCTTAAGTCCAAATTA 212
QY 768 AAACCATTTGCTAATGATGAGGCGGATGTCACATAAATCTGAGAAAGCTGACTACTTA 827
    |||||
DB 213 AAACCATTTGCTAATGATGAGGCGGATGTCACATAAATCTGAGAAAGCTGACTACTTA 272

```

QY 828 AAGCTATTCAAGAACAATTGATGCTTAACGCTCCACAGTAA CGA GCACTTCTTGAGGTC 887
 DB 273 AAGGCTATTCAAGAACAATTGATGCTTAACGCTCCACAGTAA CGA GCACTTCTTGAGGTC 332
 QY 888 ATTGATTTTGCAGGCGGATGCAACCACTTAATGATGGAACCGGCAAGGTTCTTCTTGTCGAC 947
 DB 333 ATTGATTTTGCAGGCGGATGCAACCACTTAATGATGGAACCGGCAAGGTTCTTCTTGTCGAC 392
 QY 948 AAAGATGTTCCGGTAACTTGGCCGACCCGACCTGTCGAAGAAATTTTGGCTAAGGGAAT 1007
 DB 393 AAAGATGTTCCGGTAACTTGGCCGACCCGACCTGTCGAAGAAATTTTGGCTAAGGGAAT 452
 QY 1008 GTGCGGTTAGACCATTAATAAGAAAAACAATAAAAAACAAGGAAATCTGTTGATGTG 1067
 DB 453 GTGCGGTTAGACCATTAATAAGAAAAACAATAAAAAACAAGGAAATCTGTTGATGTG 512
 QY 1068 GAATATATCTGTACAGTTTACTCCCTTAACCTGATGACGATTTTCAACAGGTTCTCAA 1127
 DB 513 GAATATATCTGTACAGTTTACTCCCTTAACCTGATGACGATTTTCAACAGGTTCTCAA 572
 QY 1128 GATACCTAAGCTATTGAAAAACATAGCTATGCGTGAACCAACATCAATCTCAAGAAATTA 1187
 DB 573 GATACCTAAGCTATTGAAAAACATAGCTATGCGTGAACCAACATCAATCTCAAGAAATTA 632
 QY 1188 GCTCAAGCAAAAGCATTTTAAACAAAAACAACCGGCTATACGATTTATGAAGCGTAC 1247
 DB 633 GCTCAAGCAAAAGCATTTTAAACAAAAACAACCGGCTATACGATTTATGAAGCGTAC 692
 QY 1248 TCCCTCAATCGTCACTCAGCAATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 1307
 DB 693 TCCCTCAATCGTCACTCAGCAATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 752
 QY 1308 TTACTCTACCGTGTAAATAATCGGGAACAAGCTTATAGATCAATAAAAAATCTGCTTG 1367
 DB 753 TTACTCTACCGTGTAAATAATCGGGAACAAGCTTATAGATCAATAAAAAATCTGCTTG 812
 QY 1368 AATGAAGAAATAAACACACATGACCTGATCTGAGAAATTTAGTGTCTTAAAAAAGG 1427
 DB 813 AATGAAGAAATAAACACACATGACCTGATCTGAGAAATTTAGTGTCTTAAAAAAGG 872
 QY 1428 GAAAGCCGATGATCCCTTTGATCGCAGTCACTTGAACCTGTTACCATCAAAATAGCTT 1487
 DB 873 GAAAGCCGATGATCCCTTTGATCGCAGTCACTTGAACCTGTTACCATCAAAATAGCTT 932
 QY 1488 GATGTGATATCCAAACGAATGCTTAAAAAGTGAAGCACTCTTAAACAGTACGGAACGTAC 1547
 DB 933 GATGTGATATCCAAACGAATGCTTAAAAAGTGAAGCACTCTTAAACAGTACGGAACGTAC 992
 QY 1548 TTAGACTTCAGAGATTTATAGATCCCTGATGATGAGCTTAACTCTACAAACATCTC 1607
 DB 993 TTAGACTTCAGAGATTTATAGATCCCTGATGATGAGCTTAACTCTACAAACATCTC 1052
 QY 1608 GATGCTTTGGTATTTATGACTATACCTTAACCTGGAAGAGTGAAGATACAGATGAC 1667
 DB 1053 GATGCTTTGGTATTTATGACTATACCTTAACCTGGAAGAGTGAAGATACAGATGAC 1112
 QY 1668 ACCAACCGTATCAATAACCGTTTATATGAGGCAAGGACCCGAAGAGAAATGCTAGCTAC 1727
 DB 1113 ACCAACCGTATCAATAACCGTTTATATGAGGCAAGGACCCGAAGAGAAATGCTAGCTAC 1172
 QY 1728 CATTTAGCTGTGTGTGCGCCAGGCGCAACAATGTTGATCCCATAGCTGAGAAAGTGTGAT 1787
 DB 1173 CATTTAGCTGTGTGTGCGCCAGGCGCAACAATGTTGATCCCATAGCTGAGAAAGTGTGAT 1232
 QY 1788 CATGCTGTGGGACTTCTGATGTGTGTGAGAAACGTTGGAGAAAGCCCTTCAAGAGCTGG 1847
 DB 1233 CATGCTGTGGGACTTCTGATGTGTGTGAGAAACGTTGGAGAAAGCCCTTCAAGAGCTGG 1292
 QY 1848 ATGATGATGATTTGATCTTCCCTGAGGAGAAAGCGACGCAATCACTTGCACTTCTAGA 1907
 DB 1293 ATGATGATGATTTGATCTTCCCTGAGGAGAAAGCGACGCAATCACTTGCACTTCTAGA 1352

QY 1908 AATAGATCAACGATCAGGACACAAAGACATCTATAGAAATTGAGACACCTGAGCAAG 1967
 DB 1353 AATAGATCAACGATCAGGACACAAAGACATCTATAGAAATTGAGACACCTGAGCAAG 1412
 QY 1968 AAGATTAATCCAGAAACCTGCTCCAGTGCATCTGCAACGAGCAACGCGCCGAGAGATGCG 2027
 DB 1413 AAGATTAATCCAGAAACCTGCTCCAGTGCATCTGCAACGAGCAACGCGCCGAGAGATGCG 1472
 QY 2028 AAGTGTGAGGACACACCTCTGTGCAAGACCAATGAGGCGAATCTGGCCCTTCAACCAT 2087
 DB 1473 AAGTGTGAGGACACACCTCTGTGCAAGACCAATGAGGCGAATCTGGCCCTTCAACCAT 1532
 QY 2088 GTTCGTTAG 2096
 DB 1533 GTTCGTTAG 1541
 RESULT 4
 AAA37637
 ID AAA37637 standard; DNA; 1661 BP.
 XX
 AC AAA37637;
 XX
 DT 15-SEP-2003 (revised)
 DT 13-OCT-2000 (first entry)
 XX
 DE Chimeric SK-FBD coding sequence.
 XX
 KM Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
 KM plasminogen; human; fibrinectin; thrombolytic therapy;
 KM cardiovascular disorder; ss.
 XX
 OS Streptococcus dysgalactiae subsp. equisimilis.
 OS Homo sapiens.
 OS Chimeric.
 XX
 PN EP1024192-A2.
 XX
 PD 02-AUG-2000.
 XX
 PF 23-DEC-1999; 99EP-00310541.
 XX
 PR 24-DEC-1998; 98IN-DE003825.
 XX
 PA (COUL) CSIR COUNCIL SCT IND RES.
 XX
 PI Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
 PI Yadav M;
 XX
 DR WPI; 2000-516032/47.
 XX
 PT Hybrid streptokinase-fibrin binding domain polypeptides useful for
 PT thrombolytic therapy comprises a streptokinase fused with fibrin binding
 PT domains of human fibrinectin.
 XX
 PS Example 5; Fig 19b; 58bp; English.
 XX
 CC This sequence represents a chimeric streptokinase-fibrin binding domain
 CC (SK-FBD) protein coding sequence. The invention relates to a hybrid
 CC plasminogen activator (PA) comprises a polypeptide fusion between
 CC streptokinase (SK), which are capable of plasminogen (PG) activation, and
 CC fibrin binding regions of human fibrinectin, which are from fibrin
 CC binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA possesses the
 CC ability to bind with fibrin independently and also characteristically
 CC retains a PG activation ability which becomes evident only after a
 CC pronounced duration, or lag, after exposure of the PA to a suitable
 CC animal or human PG. The hybrid streptokinase-fibrin binding domain
 CC polypeptides are useful in thrombolytic therapy for various kinds of
 CC cardiovascular disorders. The hybrids have enhanced fibrin selectivity as
 CC well as kinetics of plasminogen activation that are distinct from that of
 CC natural streptokinase in being characterised by a temporary delay, or lag
 CC of several minutes in the natural rate of the catalytic conversion of
 CC plasminogen to plasmin (i.e. delayed-action thrombolytic). The proteins

CC can bind tightly with fibrin in blood clots soon after introduction into
 CC the vascular system without significantly activating the circulating
 CC blood plasminogen to plasmin, thus aiding in the localisation of the
 CC plasminogen activation process to the site of pathological thrombus. This
 CC overcomes systemic plasminogen activation encountered during clinical use
 CC of streptokinase. (Updated on 15-SEP-2003 to standardise OS field)

XX Sequence 1661 BP; 529 A; 370 C; 357 G; 405 T; 0 U; 0 Other;

Query Match 55.6%; Score 1165.8; DB 3; Length 1661;
 Best Local Similarity 99.4%; Pred. No. 0;
 Matches 1170; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

OY 588 ATTGCTGACCTGAGTGGCTGCTAGACCGTCACTGTGCAACACAGCAATGGTGT 647
    |||||
DB 183 ATTGCTGACCTGAGTGGCTGCTAGACCGTCACTGTGCAACACAGCAATGGTGT 242
OY 648 AGCGTTGCTGTACTGTGAGGGGACGATCAAGACATTAAGTCTTAAATTTTGAATC 707
    |||||
DB 243 AGCGTTGCTGTACTGTGAGGGGACGATCAAGACATTAAGTCTTAAATTTTGAATC 302
OY 708 GATCTAACATCAGACCTGCTCATGAGAAAGACAGCAAGCTTAACTCAAAATCA 767
    |||||
DB 303 GATCTAACATCAGACCTGCTCATGAGAAAGACAGCAAGCTTAACTCAAAATCA 362
OY 768 AAACCATTTGCTAGTAGTGGGCGCATGTCATTAACCTTGAAGAAAGCTTACTTA 827
    |||||
DB 363 AAACCATTTGCTAGTAGTGGGCGCATGTCATTAACCTTGAAGAAAGCTTACTTA 422
OY 828 AAGGCTATTCAGAAACATTAATGCTTAAGCTCCACAGTAACGACACTACTTGAAGTC 887
    |||||
DB 423 AAGGCTATTCAGAAACATTAATGCTTAAGCTCCACAGTAACGACACTACTTGAAGTC 482
OY 888 ATTGATTTTTCAGAACGATGCAACCATTACTGATGAAAACGGCAAGTCTACTTGTGAC 947
    |||||
DB 483 ATTGATTTTTCAGAACGATGCAACCATTACTGATGAAAACGGCAAGTCTACTTGTGAC 542
OY 948 AAAGATGTTGGGTAACCTTGGCCGACCCCAACCTGTCCAAAGATTTTGTCTAAGGGACAT 1007
    |||||
DB 543 AAAGATGTTGGGTAACCTTGGCCGACCCCAACCTGTCCAAAGATTTTGTCTAAGGGACAT 602
OY 1008 GTGCGCGTTAGACCATTAAGAAACCAATACAAACCAAGCAATCTGTGATGTG 1067
    |||||
DB 603 GTGCGCGTTAGACCATTAAGAAACCAATACAAACCAAGCAATCTGTGATGTG 662
OY 1068 GAATATACCTGTACAGTTTACTCCCTTAAACCTGTGATGAGATTTCAAGACAGTCTCAA 1127
    |||||
DB 663 GAATATACCTGTACAGTTTACTCCCTTAAACCTGTGATGAGATTTCAAGACAGTCTCAA 722
OY 1128 GATACCTAAGCTATTGAAAAACATGCTATGCTGACACCATCATCTCAAGATTTACTA 1187
    |||||
DB 723 GATACCTAAGCTATTGAAAAACATGCTATGCTGACACCATCATCTCAAGATTTACTA 782
OY 1188 GCTCAAGCAAAAGCATTTTAAACAAAACACCCAGGCTATAGATTTTGAAGCGTAC 1247
    |||||
DB 783 GCTCAAGCAAAAGCATTTTAAACAAAACACCCAGGCTATAGATTTTGAAGCGTAC 842
OY 1248 TCCTCAATGCTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 1307
    |||||
DB 843 TCCTCAATGCTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 902
OY 1308 TTACTCTTACCGTGTAAATAATCGGGAACAGCTTATAGATCAATTAATAATCTGTGTG 1367
    |||||
DB 903 TTACTCTTACCGTGTAAATAATCGGGAACAGCTTATAGATCAATTAATAATCTGTGTG 962
OY 1368 AATGAAGAAATAAACAACACGACCTGATCTGAGAAATTTTACGTCCTTAAAAAGGG 1427
    |||||
DB 963 AATGAAGAAATAAACAACACGACCTGATCTGAGAAATTTTACGTCCTTAAAAAGGG 1022
OY 1428 GAAAGACCGTATGATCCCTTTGATGCGAGTCACTTGAACCTGTTCACCATCAAAATACGTT 1487
    |||||
DB 1023 GAAAGACCGTATGATCCCTTTGATGCGAGTCACTTGAACCTGTTCACCATCAAAATACGTT 1082

```

```

OY 1488 GATGTCGATPACCAAGAAATGCTTAAAGAAAGTACGAGCTTAAACAGTACGAGATAC 1547
    |||||
DB 1083 GATGTCGATPACCAAGAAATGCTTAAAGAAAGTACGAGCTTAAACAGTACGAGATAC 1142
OY 1548 TTAGACTTCAGAGATTTTAAAGATCTCTGTATGATTAAGCTTAACTTACCAACATCTC 1607
    |||||
DB 1143 TTAGACTTCAGAGATTTTAAAGATCTCTGTATGATTAAGCTTAACTTACCAACATCTC 1202
OY 1608 GATGCTTTTGTATTAATGAGTATPACCTTAACCTGGAAGAAAGTAAAGATTAATACATGAC 1667
    |||||
DB 1203 GATGCTTTTGTATTAATGAGTATPACCTTAACCTGGAAGAAAGTAAAGATTAATACATGAC 1262
OY 1668 ACCAACCGTATCATPACCGTTTAAATGAGGCAAGCCGAAAGAGAGAAATGCTATGCTAC 1727
    |||||
DB 1263 ACCAACCGTATCATPACCGTTTAAATGAGGCAAGCCGAAAGAGAGAAATGCTATGCTAC 1322
OY 1728 CATTTAGCTGTGTGTGGCCAGGCGCAACAGATTGTAC 1764
    |||||
DB 1323 CATTTAGCTGTGTGTGTGGCCAGGCGCAACAGATTGTTC 1359

```

RESULT 5

```

AAX80497
ID AAX80497 standard; cDNA; 2385 BP.
XX
XX AAX80497;
XX
XX 26-AUG-1999 (first entry)
DT
XX
XX Streptokinase and maltose binding protein fusion protein encoding cDNA.
DE
XX Streptococcus; streptokinase; fibrin-dependent plasminogen activator;
KM nSK; Bacterial; blood clot; thrombotic condition;
KM myocardial infarction; venous thrombosis; pulmonary embolism;
KM cerebral thrombosis; graft thrombosis; arterial thrombosis; ss.
XX
OS Streptococcus dysgalactiae subsp. equisimilis.
OS Synthetic.
XX
XX WO9331247-A1.
XX
XX 24-JUN-1999.
XX
XX 15-DEC-1998; 98WO-US026694.
XX
XX 15-DEC-1997; 97US-0069497P.
XX
XX (HARD ) HARVARD COLLEGE.
XX
XX Reed GU;
XX
XX WPI; 1999-395183/33.
XX
XX P-PSDB; AAY24797.
XX
XX N-terminally deleted streptokinase.
XX
XX Example; Page 45-48; 73bp; English.
XX
XX The present invention describes an isolated bacterial protein that
CC induces fibrin-dependent plasminogen activation in a pharmaceutical
CC composition for dissolving blood clots. Also described are: (1) a
CC composition comprising an isolated modified streptokinase; the
CC modification being removal of amino acid residues in the amino terminus;
CC (2) a method for dissolving a blood clot in a subject, comprising
CC administering to the subject a fibrin-dependent streptokinase protein; a
CC nucleic acid (I) encoding a modified bacterial streptokinase; (3) an
CC expression vector comprising (I); and (4) a host cell transformed with
CC the expression vector of (3). The pharmaceutical composition comprising a
CC bacterial fibrin-dependent plasminogen activator is useful for dissolving
CC blood clots in patients with a thrombotic condition, e.g. myocardial
CC infarction, venous thrombosis, pulmonary embolism, cerebral thrombosis,
CC graft thrombosis and arterial thrombosis. The modified streptokinase can
CC also be used in non-human mammals. Streptokinase activation of

```

CC plasminogen is at least 10-fold, preferably 100-fold greater in the
 CC presence of fibrin than in the absence of fibrin. The modified
 CC streptokinase has at least one amino acid substitution that inactivates a
 CC substrate site for proteolytic cleavage. This reduces the rate of
 CC degradation of the streptokinase at least two-fold. The present sequence
 CC encodes a streptokinase and maltose binding protein fusion protein from
 CC an example of the present invention

XX Sequence 2385 BP; 745 A; 547 C; 549 G; 544 T; 0 U; 0 Other;

Query Match 54.9%; Score 1150.8; DB 2; Length 2385;

Best Local Similarity 99.0%; Pred. No. 0; Mismatches 12; Indels 0; Gaps 0;

Db 573 TTCACGATGTTCTATATTCGACCTGAGCGCTGTAGACCCGTCATCTGTCAACAC 632
 1129 TCCATCGAGGGTAGATTCCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAC 1188
 QY 633 AGCCAAATGGTTGTTAGCGTTGCTGTTACTGTGAGGGAGCAATCAAGACATTAGCTT 692
 1189 AGCCAAATGATGTTAGTGGTTGCTGTTACTGTGAGGGAGCAATCAAGACATTAGCTT 1248
 QY 693 AAATTTTGAATTCGATCTAACATCAAGACCTGCTCATGAGGAAAGACAGAGCAAGC 752
 1249 AAATTTTGAATTCGATCTAACATCAAGACCTGCTCATGAGGAAAGACAGAGCAAGC 1308
 QY 753 TTAAGTCAAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 812
 1309 TTAAGTCAAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 1368
 QY 813 AAAGTCACTTACTTAAAGGCTATTCAGAAATGATCGTTAAAGTCACTTAAAGGCT 872
 1369 AAAGTCACTTACTTAAAGGCTATTCAGAAATGATCGTTAAAGTCACTTAAAGGCT 1428
 QY 873 GACTACTTGAAGCTCATGATTTTTCAGAGGATGCAACCTTACTGATCGAAACGGCAAG 932
 1429 GACTACTTGAAGCTCATGATTTTTCAGAGGATGCAACCTTACTGATCGAAACGGCAAG 1488
 QY 933 GTTACTTGTGCTGACAAAGATGTTGGGTAACTTGGCCGACCCCACTGTCCAAAGATT 992
 1489 GTTACTTGTGCTGACAAAGATGTTGGGTAACTTGGCCGACCCCACTGTCCAAAGATT 1548
 QY 993 TTGCTAAGCGGACATGTCGGGTTAGACATATTAAGAAACCAATCAAAACCAACG 1052
 1549 TTGCTAAGCGGACATGTCGGGTTAGACATATTAAGAAACCAATCAAAACCAACG 1608
 QY 1053 AAATCTGTTGATGGAATATCTGTAAGTTTACTCCCTTAAACCTGATGACGATTTC 1112
 1609 AAATCTGTTGATGGAATATCTGTAAGTTTACTCCCTTAAACCTGATGACGATTTC 1668
 QY 1113 AGACCAAGGTTCAAGATTAAGCTATTTGAAAACATGACTATCGGTGACACCATCA 1172
 1669 AGACCAAGGTTCAAGATTAAGCTATTTGAAAACATGACTATCGGTGACACCATCA 1728
 QY 1173 TCTCAAGAATTACTAGTCAAGCACAAGCAATTTTAAACAAACCAACCCGCGTATCG 1232
 1729 TCTCAAGAATTACTAGTCAAGCACAAGCAATTTTAAACAAACCAACCCGCGTATCG 1788
 QY 1233 ATTATGAAGCGTACTCTCAATGTCATGACATGACATGATTTTCCGTACGATTTTA 1292
 1789 ATTATGAAGCGTACTCTCAATGTCATGACATGACATGATTTTCCGTACGATTTTA 1848
 QY 1293 CCAATGATCAAGAGTTTACTTACCGTGTAAATCGGAAACAAAGCTTATAGATCAAT 1352
 1849 CCAATGATCAAGAGTTTACTTACCGTGTAAATCGGAAACAAAGCTTATAGATCAAT 1908
 QY 1353 AAAAATCTGCTGATGAAGAAATTAACAACACCTGACCTGATCTGAGAAATATTAC 1412
 1909 AAAAATCTGCTGATGAAGAAATTAACAACACCTGACCTGATCTGAGAAATATTAC 1968
 QY 1413 GTCTTAAAAAGGGGAAAGCGGTATGATCCCTTGAATGCGACGCTTGAACCTGTT 1472
 1969 GTCTTAAAAAGGGGAAAGCGGTATGATCCCTTGAATGCGACGCTTGAACCTGTT 2028

QY 1473 ACCATCAATATCGTTGATGTCATACCAAGATTCCTAAAAAGTAGAGCTCTTAA 1532
 2029 ACCATCAATATCGTTGATGTCATACCAAGATTCCTAAAAAGTAGAGCTCTTAA 2088
 QY 1533 GCTAGCGAAAGTAACTTGAAGCTTCAAGATTTATATGATCTCTGATTAAGGCTAA 1592
 2089 GCTAGCGAAAGTAACTTGAAGCTTCAAGATTTATATGATCTCTGATTAAGGCTAA 2148
 QY 1593 CTCTCAACAATCTGATGCTTTTGTATATGATCTTAACTTAACTGAAAGTAGAG 1652
 2149 CTCTCAACAATCTGATGCTTTTGTATATGATCTTAACTGAAAGTAGAG 2208
 QY 1653 GATAATCAAGATGACACCAACCGTATCATPACCGTTTATATGAGCAAGGACCCGAAG 1712
 2209 GATAATCAAGATGACACCAACCGTATCATPACCGTTTATATGAGCAAGGACCCGAAG 2268
 QY 1713 GAGAAATGCTAGCTTACATTTAGCTGATGCT 1742
 2269 GAGAAATGCTAGCTTACATTTAGCTGATGCT 2298
 Db

RESULT 6
 AAA37622
 ID AAA37622 standard; DNA; 1377 BP.

XX AAA37622;

AC 15-SEP-2003 (revised)
 DT 13-OCT-2000 (first entry)

XX Streptokinase-NTN gene.

XX Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
 KW plasminogen; human; fibronectin; thrombolytic therapy;
 KW cardiovascular disorder; ss.

OS Streptococcus dysgalactiae subsp. equisimilis.

FN EP1024192-A2.

PD 02-AUG-2000.

XX 23-DEC-1999; 99BP-00310541.

XX 24-DEC-1998; 98IN-DE003825.

XX (COUL) CSIR COUNCIL SCI IND RES.

PI Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
 PI Yadav M;

DR WPI; 2000-516032/47.

PT Hybrid streptokinase-fibrin binding domain polypeptides useful for
 PT thrombolytic therapy comprises a streptokinase fused with fibrin binding
 PT domain of human fibronectin.

PS Example 1; Fig 11; 58pp; English.

CC This sequence represents a streptokinase-NTN (SK-NTN) gene (where NTN
 CC stands for N-terminally repaired with native sequence). The invention
 CC relates to a hybrid plasminogen activator (PA) comprises a polypeptide
 CC fusion between streptokinase (SK), which are capable of plasminogen (PG)
 CC activation, and fibrin binding regions of human fibronectin, which are
 CC from fibrin binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA
 CC possesses the ability to bind with fibrin independently and also
 CC characteristically retains a PG activation ability which becomes evident
 CC only after a pronounced duration, or 1ag, after exposure of the PA to a
 CC suitable animal or human PG. The hybrid streptokinase-fibrin binding
 CC domain polypeptides are useful in thrombolytic therapy for various kinds
 CC of cardiovascular disorders. The hybrids have enhanced fibrin selectivity
 CC as well as kinetics of plasminogen activation that are distinct from that

CC of natural streptokinase in being characterised by a temporary delay, or
 CC lag of several minutes in the natural rate of the catalytic conversion of
 CC plasminogen to plasmin (i.e. delayed-action thrombolytic). The proteins
 CC can bind tightly with fibrin in blood clots soon after introduction into
 CC the vascular system without significantly activating the circulating
 CC blood plasminogen to plasmin, thus aiding in the localisation of the
 CC plasminogen activation process to the site of pathological thrombus. This
 CC overcomes systemic plasminogen activation encountered during clinical use
 CC of streptokinase. (Updated on 15-SEP-2003 to standardise OS field)

XX Sequence 1377 BP; 465 A; 302 C; 265 G; 345 T; 0 U; 0 Other;

Query Match 54.7%; Score 1147; DB 3; Length 1377;

Best Local Similarity 99.6%; Pred. No. 0;

Matches 1150; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

QY 588 ATTGCTGACCTGAGTGGCTGCTAGACCGTTCATCTGTCAACAAACAGCAATTGGTTT 647
DB 133 ATTGCTGACCTGAGTGGCTGCTAGACCGTTCATCTGTCAACAAACAGCAATTGGTTT 192
QY 648 AGCGTTCGCTGCTGCTGAGGGGAGCGAATCAAGCACTTAAGCTTAATTTTGAATC 707
DB 193 AGCGTTCGCTGCTGCTGAGGGGAGCGAATCAAGCACTTAAGCTTAATTTTGAATC 252
QY 708 GATCTAATCAGACGCTGCTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 767
DB 253 GATCTAATCAGACGCTGCTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 312
QY 768 AAACCAATTTGCTACTGATAGTGGCGCATGTCATATTAACCTTGAGAAAGCTGACTTA 827
DB 313 AAACCAATTTGCTACTGATAGTGGCGCATGTCATATTAACCTTGAGAAAGCTGACTTA 372
QY 828 AAGGCTATTCAAAACAAATGATGCTGCTAGCTCAAGCTCAAGCACTTCTTGAGAGTC 887
DB 373 AAGGCTATTCAAAACAAATGATGCTGCTAGCTCAAGCTCAAGCACTTCTTGAGAGTC 432
QY 888 ATTGATTTTTCAGAGCGATGCAACATCTGATGCAAGAGCAAGGCTTCTTGCTGAC 947
DB 433 ATTGATTTTTCAGAGCGATGCAACATCTGATGCAAGAGCAAGGCTTCTTGCTGAC 492
QY 948 AAAGATGCTTCGCTGTAACCTTGCAGACCCAACTGTCCAGAAATTTTGTAAAGCGGACAT 1007
DB 493 AAAGATGCTTCGCTGTAACCTTGCAGACCCAACTGTCCAGAAATTTTGTAAAGCGGACAT 552
QY 1008 GTGCGCGTTGACCATATTAAGAAACCAATACAAACCAAGCAATCTGTTGATGTG 1067
DB 553 GTGCGCGTTGACCATATTAAGAAACCAATACAAACCAAGCAATCTGTTGATGTG 612
QY 1068 GAATATACGTGATAGTTTACTCCCTTAAACCTGATGACGATTTCAAGAGGCTCAAA 1127
DB 613 GAATATACGTGATAGTTTACTCCCTTAAACCTGATGACGATTTCAAGAGGCTCAAA 672
QY 1128 GATCTAAGCTATTTGAAAACATAGCTATCGGTGACACCATCATCTCAAGAAATTACTA 1187
DB 673 GATCTAAGCTATTTGAAAACATAGCTATCGGTGACACCATCATCTCAAGAAATTACTA 732
QY 1188 GCTCAAGCACAAACATTTTAAACAAAACACCCAGGCTATAGATTTTAAAGAGTGCAC 1247
DB 733 GCTCAAGCACAAACATTTTAAACAAAACACCCAGGCTATAGATTTTAAAGAGTGCAC 792
QY 1248 TCCTCAATGCTCATGATGACATTTTCCGTACGATTTTAAACCAATGAGATCAAGAG 1307
DB 793 TCCTCAATGCTCATGATGACATTTTCCGTACGATTTTAAACCAATGAGATCAAGAG 852
QY 1308 TTTACTTACCGTGTAAATAATCGGGAAACAGCTTATAGATCAATAAATAATCTGCTCG 1367
DB 853 TTTACTTACCGTGTAAATAATCGGGAAACAGCTTATAGATCAATAAATAATCTGCTCG 912
QY 1368 AATGAAGAATAAACAACACTGACCTGATCTGAGAAATATTAACGTCCTTAAATAAGAG 1427
DB 913 AATGAAGAATAAACAACACTGACCTGATCTGAGAAATATTAACGTCCTTAAATAAGAG 972
QY 1428 GAAAGACCGATATGATCCCTTTGATCGAGTCACTTGAACCTGTCACCAATCAATACGTT 1487

```

```

DB 973 GAAAGCCGTATGATCCCTTTGATCGAGTCACTTGAACCTGTCACCATCAATACGTT 1032
QY 1488 GATTCGATATCAACAGAAATTCCTTAAAGTAGAGAGAGCTTTAAACAGCTAGAGAGTAA 1547
DB 1033 GATTCGATATCAACAGAAATTCCTTAAAGTAGAGAGAGCTTTAAACAGCTAGAGAGTAA 1092
QY 1548 TTAGACTTCAGAGATTTATATAGATCCCTGATATAGAGTAAAGTACTTACAAACAATCTC 1607
DB 1093 TTAGACTTCAGAGATTTATATAGATCCCTGATATAGAGTAAAGTACTTACAAACAATCTC 1152
QY 1608 GATCCTTTGGTATATATAGACTATACCTTAACTGAAAAGTAGAGATATACAGATGAC 1667
DB 1153 GATCCTTTGGTATATATAGACTATACCTTAACTGAAAAGTAGAGATATACAGATGAC 1212
QY 1668 ACCAACCGTATCATACCGTTTATATAGGCGAAGGAGCCGGAAGAGAGAGATGCTAGTAC 1727
DB 1213 ACCAACCGTATCATACCGTTTATATAGGCGAAGGAGCCGGAAGAGAGATGCTAGTAT 1272
QY 1728 CATTAGCTGCTGCT 1742
DB 1273 CATTAGCTATATGAT 1287

RESULT 7
AAK80492
ID AAK80492 standard; cDNA; 1242 BP.
XX
AC AAK80492;
XX
DT 17-OCT-2003 (revised)
DT 26-AUG-1999 (first entry)
XX
DE Streptococcus equisimilis native streptokinase encoding cDNA.
XX
KM Streptococcus; streptokinase; fibrin-dependent plasminogen activator;
KM nsk; RSK; bacterial; blood clot; thrombotic condition;
KM myocardial infarction; venous thrombosis; pulmonary embolism;
KM cerebral thrombosis; graft thrombosis; arterial thrombosis; ss.
XX
OS Streptococcus dysgalactiae subsp. equisimilis.
XX
PN M09931247-A1.
XX
PD 24-JUN-1999.
XX
PF 15-DEC-1998; 98WO-US026694.
XX
PR 15-DEC-1997; 97US-0069497P.
XX
PA (HARD ) HARVARD COLLEGE.
XX
PI Reed GL;
XX
DR WPI; 1999-395183/33.
XX
DR P-PsDB; AAY24794.
XX
N-terminally deleted streptokinase.
XX
PS Claim 44; Page 58-60; 73pp; English.
XX
CC The present invention describes an isolated bacterial protein that
CC induces fibrin-dependent plasminogen activation in a pharmaceutical
CC composition for dissolving blood clots. Also described are: (1) a
CC composition comprising an isolated modified streptokinase, the
CC modification being removal of amino acid residues in the amino terminus;
CC (2) a method for dissolving a blood clot in a subject, comprising
CC administering to the subject a fibrin-dependent streptokinase protein; a
CC nucleic acid (I) encoding a modified bacterial streptokinase; (3) an
CC expression vector comprising (I); and (4) a host cell transformed with
CC the expression vector of (3). The pharmaceutical composition comprising a
CC bacterial fibrin-dependent plasminogen activator is useful for dissolving
CC blood clots in patients with a thrombotic condition, e.g. myocardial

```


infection, venous thrombosis, pulmonary embolism, cerebral thrombosis, graft thrombosis and arterial thrombosis. The modified streptokinase can also be used in non-human mammals. Streptokinase activation of plasminogen is at least 10-fold, preferably 100-fold greater in the presence of fibrin than in the absence of fibrin. The modified streptokinase has at least one amino acid substitution that inactivates a substrate site for proteolytic cleavage. This reduces the rate of degradation of the streptokinase at least two-fold. The present sequence encodes native streptokinase (nsK). (Updated on 17-OCT-2003 to standardise OS field)

Sequence 1242 BP; 424 A; 267 C; 237 G; 314 T; 0 U; 0 Other;

Query Match	54.6%;	Score 1145.4;	DB 2;	Length 1242;
Best Local Similarity	99.5%;	Pred No 0.		

Matches 1149; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 588 ATGCTGGACCTCAGTGGCTGTGACACCGTCCATCTGTCAACAACAGCCCAATTGGTGT 64
 Db 1 ATTGCTGGACCTCAGTGGCTGTGACACCGTCCATCTGTCAACAACAGCCCAATTAGTGT 60

648 AGCGTTCCTGTAATCTTTGAGGGGACCAATCAAGACATTAGCTTTAAATTTTGAATC 70
 61 AGCGTTCCTGTAATCTTTGAGGGGACCAATCAAGACATTAGCTTTAAATTTTGAATC 122

Oy	708	GATCAACATACGAGCTGTCTATGGAGAAAGCAGAGCAAGGCTTAATGCCAAATCA	76
Db	121	GATTCACATACGAGCTGTCTATGGAGAAAGCAGAGCAAGGCTTAATCCAAAATCA	18

Qy	768	181	Db
AAACATTTGCTAGTAGTGGCGGATGTACATAAACTTGGAAAGCTGACTACTA	82		
AAACATTTGCTAGTAGTGGCGGATGTACATAAACTTGGAAAGCTGACTACTA	24		

Qy	828	AAGGCTATTGAGAA	CAATTGATCGTAA	CGTCCA	CAGTAA	CGACGACTA	CTTTGA	GCTC	88
Db	241	AAGGCTATTGAGAA	CAATTGATCGTAA	CGTCCA	CAGTAA	CGACGACTA	CTTTGA	GCTC	30

Qy	888	ATTGATTTTGCAGCCATGCACCATTACTGATGMAACGGCAAGTCTACTTTGCTGAC	94
Db	301	ATTGATTTTGCAGCCATGCACCATTACTGATGMAACGGCAAGTCTACTTTGCTGAC	36

QY 948 AAAGATGTTCCGTAACCTTGGCCGACCCAACTGTCCAGAAATTTTGTGAAGCGGACAT 10

Db 361 AAAGATGTTCCGTAACCTTGGCCGACCCAACTGTCCAGAAATTTTGTGAAGCGGACAT 422

QY 1008 GTGCGCGTTAGACCATATTAATAAGAAAAACCAATACAAACCAAGCGAATCTGTTGATGTG 10

DB 421 GTGCGCGTTAGACCATATTAATAAGAAAAACCAATACAAACCAAGCGAATCTGTTGATGTG 48

1068 GAATATACCTGACAGTTACTCCTTAAACCTGATGAGATTGAGACCAGGCTCAAA 11
 481 GAATATACCTGACAGTTACTCCTTAAACCTGATGAGATTGAGACCAGGCTCAAA 11
 DB

[illegible]

1188 GCTCAGCACAAGCATTTTAAACAAACCACCAGGCTATACGATTTATGAAGTGC 12

1248 TCCCTCAATCGTCACATGACGAATGACATTTCCGTACATTTTACCAATGATCAAGAG 13

1308 TTTACTTACCGGTGTTAAATCGGGAACAGCTTATAGATCAATAAAATCTGCTCG 13

1368 AATGAGAAATTAAACAACCTGACTGATCTTGAGAAATATTACGTCTTAATAAAGG 144

[illegible]

QY	1428	GAAAGCCGTAATGCCCTTTGATCGCAGTCACTTGAACGTTACATCAATAAATAGCTT	1487
Db	841	GAAGAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACGTTACATCAATAAATAGCTT	900
QY	1488	GATGTCGATACCAACGAATTCCTAAAAAGTAGCAGCTCTTAAACAGTAGCGAAAGCTAAC	1547
Db	901	GATGTCGATACCAACGAATTCCTAAAAAGTAGCAGCTCTTAAACAGTAGCGAAAGCTAAC	960
QY	1548	TTAAGCTTCAGAGATTATATACGATCCTCGTAGTAAGGCTAAACCTACCTCAACAACATCTC	1607
Db	961	TTAAGCTTCAGAGATTATATACGATCCTCGTAGTAAGGCTAAACCTACCTCAACAACATCTC	1020
QY	1608	GATGCTTTGGTATTTATAGACTATACCTTAACTGAAAAAGTAGAGATTAATCAAGATGAC	1667
Db	1021	GATGCTTTGGTATTTATAGACTATACCTTAACTGAAAAAGTAGAGATTAATCAAGATGAC	1080
QY	1668	ACCAACCGTATCATTAACGTTTATATGGGCAAGGCAACCCGAAAGAGAGAAATGCTATAGCTAC	1727
Db	1081	ACCAACCGTATCATTAACGTTTATATGGGCAAGGCAACCCGAAAGAGAGAAATGCTATAGCTAT	1140
QY	1728	CATTAGCTGCTGCT 1742	
Db	1141	CATTAGCTATGAT 1155	

RESULT	8
AAA37633	
ID	AAA37633 standard; DNA; 1245 BP.
XY	

AC AAA37633 ;

DT 15-SEP-2003 (revised)

2000

DE S. equisimilis streptokinase coding sequence.
TV

KW Streptokinase; SK, hybrid plasminogen activator; fibrin binding region;

KW cardiovascular disorder; ss.

05 *Streptococcus dysgalactiae* subsp. *equisimilis*.
 xv

FH	Key	Location/Qualifiers
FM	and	1 1045

/*tag= a

2000

XX

2000

2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841

XX
XX

22

PI Yadav M;

DR WPI; 2000-516032/47.

XX
XX
XX

PT thrombolytic therapy comprises a streptokinase fused with fibrin binding

XX XX

XX

CC streptokinase coding sequence. The invention relates to a hybrid

CC streptokinase (SK), which are capable of plasminogen (PG) activation, and

CC fibrin binding regions of human fibronectin, which are from fibrin
 CC binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA possesses the
 CC ability to bind with fibrin independently and also characteristically
 CC retains a pg activation ability which becomes evident only after a
 CC pronounced duration, or lag, after exposure of the PA to a suitable
 CC animal or human Pg. The hybrid streptokinase-fibrin binding domain
 CC polyepitides are useful in thrombolytic therapy for various kinds of
 CC cardiovascular disorders. The hybrids have enhanced fibrin selectivity as
 CC well as kinetics of plasminogen activation that are distinct from that of
 CC natural streptokinase in being characterised by a temporary delay, or lag
 CC of several minutes in the natural rate of the catalytic conversion of
 CC plasminogen to plasmin (i.e. delayed-action thrombolytic). The proteins
 CC can bind tightly with fibrin in blood clots soon after introduction into
 CC the vascular system without significantly activating the circulating
 CC blood plasminogen to plasmin, thus aiding in the localisation of the
 CC plasminogen activation process to the site of pathological thrombus. This
 CC overcomes systemic plasminogen activation encountered during clinical use
 CC of streptokinase. (Updated on 15-Sep-2003 to standardise OS field)

XX Sequence 1245 BP; 426 A; 267 C; 237 G; 315 T; 0 U; 0 Other;

Query Match 54.6%; Score 1145.4; DB 3; Length 1245;

Best Local Similarity 99.5%; Pred. No. 0;

Matches 1149; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 588 ATTGCTGACCTGAGTGGCTGCTAGACGTCATCTGTCAACACAGCAATGGTGT 647
 DB 1 ATTGCTGACCTGAGTGGCTGCTAGACGTCATCTGTCAACACAGCAATGGTGT 60
 QY 648 AGCGTGTCTGTAAGTGGGAGCAATCAAGACATTAAGTCTTAAATTTTGAATC 707
 DB 61 AGCGTGTCTGTAAGTGGGAGCAATCAAGACATTAAGTCTTAAATTTTGAATC 120
 QY 708 GATCTAACATCAGACCTGCTCATGAGAAACAGAGCAAGGCTTAAGTCCAAATCA 767
 DB 121 GATCTAACATCAGACCTGCTCATGAGAAACAGAGCAAGGCTTAAGTCCAAATCA 180
 QY 768 AAACCATTTGCTACTGATGTCGGGAGATGTCATTAACCTTGAGAAAGCTTA 827
 DB 181 AAACCATTTGCTACTGATGTCGGGAGATGTCATTAACCTTGAGAAAGCTTA 240
 QY 828 AAGGCTATTCAGAACATTAATGCTCAAGTCAAGTCAAGTCAAGTCAAGTCA 887
 DB 241 AAGGCTATTCAGAACATTAATGCTCAAGTCAAGTCAAGTCAAGTCAAGTCA 300
 QY 888 ATTGATTTTTCAGACGATGCAATTAAGTCAAGTCAAGTCAAGTCAAGTCA 947
 DB 301 ATTGATTTTTCAGACGATGCAATTAAGTCAAGTCAAGTCAAGTCAAGTCA 360
 QY 948 AAAGATGTCGGTGAACCTTGCAGACCACTGTCGAAGTCTTAAGGCGACAT 1007
 DB 361 AAAGATGTCGGTGAACCTTGCAGACCACTGTCGAAGTCTTAAGGCGACAT 420
 QY 1008 GTGCGGTTAGACCATTAATAAAGAAACCAATCAAAACCAAGGAATCTGTGATGTG 1067
 DB 421 GTGCGGTTAGACCATTAATAAAGAAACCAATCAAAACCAAGGAATCTGTGATGTG 480
 QY 1068 GAATATATCTGTACAGTTTACTCCCTTAAACCTGATGACATTTCAAGCAGGTCTCAA 1127
 DB 481 GAATATATCTGTACAGTTTACTCCCTTAAACCTGATGACATTTCAAGCAGGTCTCAA 540
 QY 1128 GATATCTAAGTATTTGAAAAACATAGCTATGCTGACCACTCATCTCAAGATTTACTA 1187
 DB 541 GATATCTAAGTATTTGAAAAACATAGCTATGCTGACCACTCATCTCAAGATTTACTA 600
 QY 1188 GCTCAAGACCAAGCAATTTTAAACAAACCAACCAAGGCTATATGAGTCAAGTCA 1247
 DB 601 GCTCAAGACCAAGCAATTTTAAACAAACCAACCAAGGCTATATGAGTCAAGTCA 660
 QY 1248 TCTCTCAATGCTCACTCATGACCAATTTTTCGATGCAATTTTCAACATGATCAAGAG 1307
 DB 661 TCTCTCAATGCTCACTCATGACCAATTTTTCGATGCAATTTTCAACATGATCAAGAG 720

QY 1308 TTATCTACCGGTGTTAAAAATCGGGAACACGCTTATAGATCAATAAAAATCTGCTCG 1367
 DB 721 TTATCTACCGGTGTTAAAAATCGGGAACACGCTTATAGATCAATAAAAATCTGCTCG 780
 QY 1368 AATGAGAAATTAACACACGCTGACCTGATCTCTGAGAAATTTACGTCCTTAAAAAGGG 1427
 DB 781 AATGAGAAATTAACACACGCTGACCTGATCTCTGAGAAATTTACGTCCTTAAAAAGGG 840
 QY 1428 GAAAAGCGGTATGATCCCTTGTATGCGAGTCACTTGAACGTTCCACATCAATACGTT 1487
 DB 841 GAAAAGCGGTATGATCCCTTGTATGCGAGTCACTTGAACGTTCCACATCAATACGTT 900
 QY 1488 GATGCTGATACCAAGCAATTCCTTAAAAAGTGAAGAGCTTTAAACAGTGAAGCAAGTAC 1547
 DB 901 GATGCTGATACCAAGCAATTCCTTAAAAAGTGAAGAGCTTTAAACAGTGAAGCAAGTAC 960
 QY 1548 TTAGACTTCAGAGATTTATATACGATCTCTGTATGAGTAAAGCTTAACTCTACCAATCTC 1607
 DB 961 TTAGACTTCAGAGATTTATATACGATCTCTGTATGAGTAAAGCTTAACTCTACCAATCTC 1020
 QY 1608 GATGCTTTGTTATATGAGTATACCTTAACTGGAAGATGAGATATCAAGATGAC 1667
 DB 1021 GATGCTTTGTTATATGAGTATACCTTAACTGGAAGATGAGATATCAAGATGAC 1080
 QY 1668 ACCAAGCGTATCATACCGTTTATATGAGGCAAGGCAAGGCAAGGCAAGGCAAGTCTAGTAC 1727
 DB 1081 ACCAAGCGTATCATACCGTTTATATGAGGCAAGGCAAGGCAAGGCAAGGCAAGTCTAGTAC 1140
 QY 1728 CATTAGCTGTGCT 1742
 DB 1141 CATTAGCTGTGCT 1155

RESULT 9

ABA05546
 ID ABA05546 standard; cDNA; 1254 BP.

XX ABA05546;

XX 26-FEB-2002 (first entry)

XX Streptokinase cDNA.

XX Streptokinase; cerebroprotective; cardiact; gene therapy; fusion;

XX Vasodilator; thrombolytic; angina; myocardial infarction; stroke;

XX gene therapy; maxadilan; ss.

XX Unidentified.

XX WO200185100-A2.

XX 15-NOV-2001.

XX 10-MAY-2001; 2001WO-US015209.

XX 11-MAY-2000; 2000US-00569920.

XX (GEHO) GEN HOSPITAL CORP.

XX Reddy VB, Lerner B;

XX WPI; 2002-062184/08.

XX New fusion protein or conjugate, useful for treating unstable angina,

XX acute myocardial infarction or stroke, comprises a vasodilator

XX polypeptide and a thrombolytic polypeptide, or active fragments of the

XX polypeptides.

XX Example 1; Fig 2; 37pp; English.

XX The invention relates to a fusion protein or a conjugate comprising a

XX vasodilator polypeptide, or its active fragment, and a thrombolytic

XX polypeptide or its active fragment. The protein is useful for treating a

CC subject suffering from a partially or totally occluded blood vessel,
 CC causing unstable angina, acute myocardial infarction or stroke. The
 CC polynucleotide encoding the polypeptide is useful in gene therapy. The
 CC vasodilatory action of the protein allows for the use of lower doses of a
 CC thrombolytic while maintaining the clot dissolving effectiveness of the
 CC thrombolytic, and the use of lower doses of the thrombolytic reduces
 CC associated side effects. The present sequence is the streptokinase cDNA
 CC used in the construction of a Maxadilan-Streptokinase fusion protein.
 CC Maxadilan is a vasodilator peptide produced by the salivary gland of the
 CC New World sand fly

XX Sequence 1254 BP; 426 A; 271 C; 241 G; 316 T; 0 U; 0 Other;

Query Match 54.6%; Score 1145.4; DB 6; Length 1254;

Best Local Similarity 99.5%; Pred. No. 0;

Matches 1149; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

QY 588 ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACGCAATTGGTGT 647
DB 7 ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACGCAATTGGTGT 66
QY 648 AGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 707
DB 67 AGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 126
QY 708 GATCTAACATCAGACCTGCTCATGAGAAAGACAGCAAGGCTTAAGTCAAAATCA 767
DB 127 GATCTAACATCAGACCTGCTCATGAGAAAGACAGCAAGGCTTAAGTCAAAATCA 186
QY 768 AAACCACTTGTCTAGTGTGCGCGATGTCATTAATCTTGAAGTGAAGTGAAGTGA 827
DB 187 AAACCACTTGTCTAGTGTGCGCGATGTCATTAATCTTGAAGTGAAGTGAAGTGA 246
QY 828 AAGGCTATTCAGAAACAATTGATGCTTAAGTCAAGTCAAGTCAAGTCAAGTCAAG 887
DB 247 AAGGCTATTCAGAAACAATTGATGCTTAAGTCAAGTCAAGTCAAGTCAAGTCAAG 306
QY 888 ATTGATTTTGGCAAGCAGATGCAACATTAAGTCAAGTCAAGTCAAGTCAAGTCAAG 947
DB 307 ATTGATTTTGGCAAGCAGATGCAACATTAAGTCAAGTCAAGTCAAGTCAAGTCAAG 366
QY 948 AAAGATGTTGCGTAACTTCCGACCCCACTGTCCAGAAATTTTGTGAAGCGGACAT 1007
DB 367 AAAGATGTTGCGTAACTTCCGACCCCACTGTCCAGAAATTTTGTGAAGCGGACAT 426
QY 1008 GTGCGCGTTAGACCATTAATAAGAAAAACAATAAACAAGGAAATCTGTGATGTG 1067
DB 427 GTGCGCGTTAGACCATTAATAAGAAAAACAATAAACAAGGAAATCTGTGATGTG 486
QY 1068 GAATATACCTGTACAGTTTACTCCCTTAAACCTGATGATGATTTCAACAGGCTCAAA 1127
DB 487 GAATATACCTGTACAGTTTACTCCCTTAAACCTGATGATGATTTCAACAGGCTCAAA 546
QY 1128 GATCTAACCTATTTGAACAACATAGCTATGCTGACACATCATCTCAAGAAATCTA 1187
DB 547 GATCTAACCTATTTGAACAACATAGCTATGCTGACACATCATCTCAAGAAATCTA 606
QY 1188 GCTCAAGCAAAAGCAATTTTAAACAAACACACGAGCTATAGATTTATGAAGTAC 1247
DB 607 GCTCAAGCAAAAGCAATTTTAAACAAACACACGAGCTATAGATTTATGAAGTAC 666
QY 1248 TCTCTCAATGCTCACTGATGACATGATTTTCCGATGATTTTCAAGATGATCAAG 1307
DB 667 TCTCTCAATGCTCACTGATGACATGATTTTCCGATGATTTTCAAGATGATCAAG 726
QY 1308 TTACTTACCGTGTAAATGCGGAAACAGCTTATAGATCAATTAATAATCTGCTG 1367
DB 727 TTACTTACCGTGTAAATGCGGAAACAGCTTATAGATCAATTAATAATCTGCTG 786
QY 1368 AATGAAGAATAAACAACATGATCTGTGAATAATATTAAGTCTTAAATAAGGG 1427
DB 787 AATGAAGAATAAACAACATGATCTGTGAATAATATTAAGTCTTAAATAAGGG 846

```

```

QY 1428 GAAAAGCCGTATGATCCCTTTGATGCACTGTAAGAACTGTCAACATCAATACGTT 1487
DB 847 GAAAAGCCGTATGATCCCTTTGATGCACTGTAAGAACTGTCAACATCAATACGTT 906
QY 1488 GATGCTGATACCAAGAAATTTGTAATAAAGTGAAGCTTTTAAACAGTACGCAAGTAA 1547
DB 907 GATGCTGATACCAAGAAATTTGTAATAAAGTGAAGCTTTTAAACAGTACGCAAGTAA 966
QY 1548 TTAGACTTCAGAGATTTATAGATCTGCTGTATTAAGGCTTAAACTCTACAAACAATCTC 1607
DB 967 TTAGACTTCAGAGATTTATAGATCTGCTGTATTAAGGCTTAAACTCTACAAACAATCTC 1026
QY 1608 GATGCTTTGGTATTTATGACATTAATCTTAACTGCAAGAAAGTACGATATACGATAC 1667
DB 1027 GATGCTTTGGTATTTATGACATTAATCTTAACTGCAAGAAAGTACGATATACGATAC 1086
QY 1668 ACCAACCGTATCATACCGTTTATATGCGCAAGCAACCGCAAGAGAGAAATCTGATAC 1727
DB 1087 ACCAACCGTATCATACCGTTTATATGCGCAAGCAACCGCAAGAGAGAAATCTGATAC 1146
QY 1728 CATTTAGCTGTGCT 1742
DB 1147 CATTTAGCTGTGAT 1161

```

RESULT 10

ABA05547 standard; DNA; 8893 BP.

ABA05547;

26-FEB-2002 (first entry)

Maxadilan-streptokinase fusion protein plasmid pTVB3maxsek.

Streptokinase; cerebroprotective; cardiac; gene therapy; fusion;

vasodilator; thrombolytic; angina; myocardial infarction; stroke;

gene therapy; maxadilan, sand fly; plasmid; ds.

Lutzomyia longipalpis.

OS Unidentified.

OS Synthetic.

OS Chimeric.

WO200185100-A2.

15-NOV-2001.

10-MAY-2001; 2001WO-US015209.

11-MAY-2000; 2000US-00569920.

(GENO) GEN HOSPITAL CORP.

Reddy VB, Lerner E;

WPI; 2002-062184/08.

New fusion protein or conjugate, useful for treating unstable angina,

acute myocardial infarction or stroke, comprises a vasodilator

polypeptide and a thrombolytic polypeptide, or active fragments of the

polypeptides.

Example 1; Fig 3; 37pp; English.

The invention relates to a fusion protein or a conjugate comprising a

vasodilator polypeptide, or its active fragment, and a thrombolytic

polypeptide or its active fragment. The protein is useful for treating a

subject suffering from a partially or totally occluded blood vessel,

causing unstable angina, acute myocardial infarction or stroke. The

polynucleotide encoding the polypeptide is useful in gene therapy. The

vasodilatory action of the protein allows for the use of lower doses of a

thrombolytic while maintaining the clot dissolving effectiveness of the

CC thrombolytic, and the use of lower doses of the thrombolytic reduces
CC associated side effects. The present sequence is the plasmid used for the
CC preparation of Maxadilan-Streptokinase fusion protein. It contains
CC maxadilan cDNA, which encodes a vasodilator peptide produced by the
CC salivary gland of the New World sand fly, and streptokinase cDNA

XX Sequence 8893 BP; 2366 A; 2106 C; 2246 G; 2175 T; 0 U; 0 Other;

Query Match 54.6%; Score 1145.4; DB 6; Length 8893;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1149; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 588 ATTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCAATGGTGT 647
DB 5935 ATTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCAATGGTGT 5994
QY 648 AGCGTTGCTGTGATCTGTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTGAATC 707
DB 5995 AGCGTTGCTGTGATCTGTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTGAATC 6054
QY 708 GATCTTAACATCAGACCTGCTCATGAGGAAAGCAAGCAAGGCTTAAGTCCAAATCA 767
DB 6055 GATCTTAACATCAGACCTGCTCATGAGGAAAGCAAGCAAGGCTTAAGTCCAAATCA 6114
QY 768 AAACCATTTGCTACTGATAGTGGCGCATGTCACTATAACTTGAAGAGCTGACTTA 827
DB 6115 AAACCATTTGCTACTGATAGTGGCGCATGTCACTATAACTTGAAGAGCTGACTTA 6174
QY 828 AAGGCTATTCAAGAACATTTGATCGCTAACGTCACAGTAAAGCACTACTTGAAGT 887
DB 6175 AAGGCTATTCAAGAACATTTGATCGCTAACGTCACAGTAAAGCACTACTTGAAGT 6234
QY 888 ATTGATTTTGCAGACCGCATGACCATTTACTGATGAAACGGCAAGGTCTACTTGTGAC 947
DB 6235 ATTGATTTTGCAGACCGCATGACCATTTACTGATGAAACGGCAAGGTCTACTTGTGAC 6294
QY 948 AAAGATGTTGGGTAACTTGGCGAAGCCCACTGTCCAAAGATTTTGTAAAGGAGCAT 1007
DB 6295 AAAGATGTTGGGTAACTTGGCGAAGCCCACTGTCCAAAGATTTTGTAAAGGAGCAT 6354
QY 1008 GTGCGGTTAGACCATATAAAGAAAAACAATCAAAACCAAGCAATCTGTGATGTG 1067
DB 6355 GTGCGGTTAGACCATATAAAGAAAAACAATCAAAACCAAGCAATCTGTGATGTG 6414
QY 1068 GAATATACTGTACAGTTTACTCCCTTAAACCTGATGACGATTTCAAGCAAGGTCTCAA 1127
DB 6415 GAATATACTGTACAGTTTACTCCCTTAAACCTGATGACGATTTCAAGCAAGGTCTCAA 6474
QY 1128 GATCTAAGCTATTGAAAACTAGCTATGGTGACCAATCATCTCAAGATTA 1187
DB 6475 GATCTAAGCTATTGAAAACTAGCTATGGTGACCAATCATCTCAAGATTA 6534
QY 1188 GCTCAAGCAAAAGCAATTTTAAACAAAACCAAGGCTTACGATTTAAGACGTGAC 1247
DB 6535 GCTCAAGCAAAAGCAATTTTAAACAAAACCAAGGCTTACGATTTAAGACGTGAC 6594
QY 1248 TCTCTCATGTCTCATGACCAATGACATTTTCCGATCGATTTTCAATGATCAAGAG 1307
DB 6595 TCTCTCATGTCTCATGACCAATGACATTTTCCGATCGATTTTCAATGATCAAGAG 6654
QY 1308 TTCTACTACCGTGTAAATAATCGGGAACAAGCTTATGATTAATAAATTCGTGTG 1367
DB 6655 TTCTACTACCGTGTAAATAATCGGGAACAAGCTTATGATTAATAAATTCGTGTG 6714
QY 1368 AATGAAGAATAAACAACATGACCTGATCTGTGAGAAATATTAGTCTTTAAAAAGGG 1427
DB 6715 AATGAAGAATAAACAACATGACCTGATCTGTGAGAAATATTAGTCTTTAAAAAGGG 6774
QY 1428 GAAAGCCGTATGATCTTTTGAATCGAGTCACTTGAACCTGTTCAACATCAATACGTT 1487
DB 6775 GAAAGCCGTATGATCTTTTGAATCGAGTCACTTGAACCTGTTCAACATCAATACGTT 6834
QY 1488 GATGTGATACCAACGAATTTGCTAAAAAGTAGAGAGCTCTTAACAGCTAGCGAAAGTAC 1547

DB 6835 GATGTGATACCAACGAATTTGCTAAAAAGTAGAGAGCTCTTAACAGCTAGCGAAAGTAC 6894
QY 1548 TTAAAGCTTCAAGATTTTATAGATCTCTGTGATTAAGGCTTAAACTTCTTCAACAATCTC 1607
DB 6895 TTAAAGCTTCAAGATTTTATAGATCTCTGTGATTAAGGCTTAAACTTCTTCAACAATCTC 6954
QY 1608 GATGCTTTTGGTATTATAGACTTATACCTTAACCTGGAAGAGATATACAGATGAC 1667
DB 6955 GATGCTTTTGGTATTATAGACTTATACCTTAACCTGGAAGAGATATACAGATGAC 7014
QY 1668 ACCAAGCTATCATTAACCGTTTATATGGCAAGGACCCGAAGAGAGAAATGCTAGTAC 1727
DB 7015 ACCAAGCTATCATTAACCGTTTATATGGCAAGGACCCGAAGAGAGAAATGCTAGTAC 7074
QY 1728 CATTTAGCTGTGCT 1742
DB 7075 CATTTAGCTGTATGAT 7089

RESULT 11

AAK1632 standard; DNA; 1242 BP.

AAK1632;
17-OCT-2003 (revised)
04-MAY-1999 (first entry)
Streptococcus equisimilis native streptokinase encoding DNA.

Streptococcus equisimilis H46a; streptokinase; mutant; fibrinolytic;
plasma clot; hydrolysis; haemolytic Streptococcus; plasminogen; plasmin;
serine protease; fibrin; blood clot; thrombolytic;
vascular thromboembolytic symptom; acute myocardial infarction;
fibrinolysis; resistance; ds.

Streptococcus dysgalactiae subsp. equisimilis.

Key Location/Qualifiers

1. 1242
/*tag= a
/transl_except= (pos:40..42,aa:Aen)
/note= "no stop codon given"

US5876999-A.

02-MAR-1999.

06-DEC-1995; 95US-00568393.

06-DEC-1995; 95US-00568393.

(NASC-) NAT SCI COUNCIL.

Wu H;

WPI; 1999-189643/16.

P-PSDB; AAW94664.

Mutant streptokinase polypeptide - useful as plasmin-resistant

thrombolytic agent.

Claim 1; Col 7-10; 17p; English.

The present invention describes a mutant streptokinase (SK) polypeptide
in which at least one of the amino acids in the Pro58-Lys59-Ser60-Lys61
segment of the corresponding native SK is replaced by another amino acid.
The present sequence encodes native SK. SK is a secretory protein of
haemolytic Streptococcus able to activate human plasminogen (HPIg) to
plasmin (HPIa), which is a serine protease able to catalyse the
hydrolysis of fibrin in blood clots. The SK is useful as a thrombolytic
agent in the treatment of vascular thromboembolytic symptoms such as

CC acute myocardial infarction. Compared with wild-type SK, the K59E mutant
 CC is more resistant to degradation by human plasmin and is more effective
 CC both in acting as a fibrinolytic agent and in activating human plasminogen.
 CC (Updated on 17-OCT-2003 to standardise OS field)
 CC XX

Sequence 1242 BP; 423 A; 268 C; 237 G; 314 T; 0 U; 0 Other;

Query Match 54.6%; Score 1143.8; DB 2; Length 1242;
 Best Local Similarity 99.4%; Pred. No. 0;
 Matches 1148; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

QY 588 ATTGCTGACCTGAGTGGCTGCTGTAACCGTCCATCTGTCAACAACGCCAATTGGTTG 647
DB 1 ATTCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACGCCAATTAGTGT 60
QY 648 AGCGTGGCTGCTGCTGCTGAGGGGACGAATCAAGACATTAGCTTAAATTTTGAATC 707
DB 61 AGCGTGGCTGCTGCTGCTGAGGGGACGAATCAAGACATTAGCTTAAATTTTGAATC 120
QY 708 GATCTAACATCAGCACTGCTCATGAGAGAAAGACAGCAAGCTTAACTCAAAATCA 767
DB 121 GATCTAACATCAGCACTGCTCATGAGAGAAAGACAGCAAGCTTAACTCAAAATCA 180
QY 768 AAACCATTTGCTAGTATGAGTGGCGCATGTCACTTAACTTGAAGAGCTGACTTACTA 827
DB 181 AAACCATTTGCTAGTATGAGTGGCGCATGTCACTTAACTTGAAGAGCTGACTTACTA 240
QY 828 AAGGCTATTCAAGAACATTTAGTGTCAAGCTCCAGTAAACGAGCTTAAAGTTC 887
DB 241 AAGGCTATTCAAGAACATTTAGTGTCAAGCTCCAGTAAACGAGCTTAAAGTTC 300
QY 888 ATTGATTTTTCAGAGCATGACCACTTATCTGATCGAAACGGCAAGTCTACTTCTGAC 947
DB 301 ATTGATTTTTCAGAGCATGACCACTTATCTGATCGAAACGGCAAGTCTACTTCTGAC 360
QY 948 AAAGATGTTGCTGTAACCTTCCGACCCCACTGTCCAAAGATTTTGTCAAGCGGACAT 1007
DB 361 AAAGATGTTGCTGTAACCTTCCGACCCCACTGTCCAAAGATTTTGTCAAGCGGACAT 420
QY 1008 GTGGCGCTTGAACAATTTAAAGAAACCAATACAAACCAAGCAATCTGTATATG 1067
DB 421 GTGGCGCTTGAACAATTTAAAGAAACCAATACAAACCAAGCAATCTGTATATG 480
QY 1068 GAATATATCTGACATTTACTCCCTTAAACCTGATGAGATTTTCAAGCAAGTCTCAA 1127
DB 481 GAATATATCTGACATTTACTCCCTTAAACCTGATGAGATTTTCAAGCAAGTCTCAA 540
QY 1128 GATCTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACTCAAGAAATTACTA 1187
DB 541 GATCTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACTCAAGAAATTACTA 600
QY 1188 GCTCAGACCAACAGATTTTAAACAAACCAACCAAGCTATAGCAATTATGACGTGAC 1247
DB 601 GCTCAGACCAACAGATTTTAAACAAACCAACCAAGCTATAGCAATTATGACGTGAC 660
QY 1248 TCCCTCAATCGCTCATGATGACATTTTCCGTAGATTTTACCAATGATCAAGAG 1307
DB 661 TCCCTCAATCGCTCATGATGACATTTTCCGTAGATTTTACCAATGATCAAGAG 720
QY 1308 TTTACTTACCGTGTAAATAATCGGAAACAGCTTATAGATCAATTAATAATCTGCTG 1367
DB 721 TTTACTTACCGTGTAAATAATCGGAAACAGCTTATAGATCAATTAATAATCTGCTG 780
QY 1368 AATGAGAAATTAACAACACTGACCTGATCTTGAGAAATTAAGTCTTAAATAAGGG 1427
DB 781 AATGAGAAATTAACAACACTGACCTGATCTTGAGAAATTAAGTCTTAAATAAGGG 840
QY 1428 GAAAGAGGCTATGATCTCTTTGATCGAGTCACTTGAACCTTTTCAACATCAATAGCTT 1487
DB 841 GAAAGAGGCTATGATCTCTTTGATCGAGTCACTTGAACCTTTTCAACATCAATAGCTT 900
QY 1488 GATGTCATACCAAGATTTGCTAAAGTAGAGAGGCTTAAACGATGAGGAAGCTTAC 1547

```

```

DB 901 GATGTCATACCAAGATTTGCTAAAGTAGAGAGGCTTAAACGATGAGGAAGCTTAC 960
QY 1548 TTAGACTTCAAGATTTTATACGATCTCCGATGATAGGCTTAACTTACCAACATCTC 1607
DB 961 TTAGACTTCAAGATTTTATACGATCTCCGATGATAGGCTTAACTTACCAACATCTC 1020
QY 1608 GATGCTTTTGTATTTATGACATTAACCTTAACTGAGAAAGTAGAGATTAATCAGATGAC 1667
DB 1021 GATGCTTTTGTATTTATGACATTAACCTTAACTGAGAAAGTAGAGATTAATCAGATGAC 1080
QY 1668 ACCAACCCTATCAACCGCTTTATATGAGGCAACGACCCGAGAGAGAAATGCTTAC 1727
DB 1081 ACCAACCCTATCAACCGCTTTATATGAGGCAACGACCCGAGAGAGAAATGCTTAC 1140
QY 1728 CATTTAGCTGTGGT 1742
DB 1141 CATTTAGCTGTATGAT 1155

```

```

RESULT 12
AAQ11651
ID AAQ11651 standard; DNA; 2030 BP.
XX
AC AAQ11651;
XX
AC 08-JUL-1991 (first entry)
XX
DB FB-FB-SK fusion construct.
XX
KW Fibrin-binding protein; fibrinolysis; intravascular thrombi; fibrinogen;
XX streptokinase; fusion protein; ss.
XX
OS Staphylococcus aureus.
XX
FH Key Location/Qualifiers
FT CDS 10..184
FT /*tag= a
FT /*label= FB monomer
FT CDS 185..358
FT /*tag= b
FT /*label= FB monomer
FT CDS 359..1601
FT /*tag= c
FT /*label= streptokinase
XX
FN US5011686-A.
XX
PD 30-APR-1991.
XX
PP 15-NOV-1989; 89US-00437769.
XX
PR 21-SEP-1987; 87US-00099242.
XX
PA (CREA-) CREATIVE BIOMOLEC.
XX
PI Pang RHL;
XX
XX WPI; 1991-140198/19.
XX P-PSDB; AAR11829.
XX
PT Imparting injectable fibrinolytic agent - with affinity for intravascular
XX thrombus, by linking agent to fibrin binding domain.
XX
XX Disclosure; Fig 5; 18pp; English.
XX
XX The DNA encodes an FB-FB dimer linked to the streptokinase coding
XX sequence. The FB fragment has selective affinity for fibrin, low affinity
XX for fibrinogen, and minimal immunogenicity, imparting thrombus-targeting
XX capability. See also AAQ11649 and AAQ11650
XX
SQ Sequence 2030 BP; 673 A; 464 C; 406 G; 487 T; 0 U; 0 Other;

```

Query Match 54.6%; Score 1143.8; DB 2; Length 2030;

Best Local Similarity 99.4%; Pred. No. 0;
Matches 1148; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

QY 588 ATTCGCGAAGCTGAGGCGCTGAGACCGTCCATCTCTCAACAACACGCAATGTTGTT 647
Db 358 ATTCGCGAAGCTGAGGCGCTGAGACCGTCCATCTCTCAACAACACGCAATGTTGTT 417
QY 648 AGCGTTGCTGCTGCTGAGGCGGACGATCAAGACATTAAGTCTTAATTTTGAATC 707
Db 418 AGCGTTGCTGCTGCTGAGGCGGACGATCAAGACATTAAGTCTTAATTTTGAATC 477
QY 708 GATCTAACATCAACGACCTGCTCATGAGAAAGACAGACAAAGCTTAAGTCCAAAATCA 767
Db 478 GATCTAACATCAACGACCTGCTCATGAGAAAGACAGACAAAGCTTAAGTCCAAAATCA 537
QY 768 AAACCAATTTGCTATGATAGTGGCGCATGTCACTAACTTGAGAAAGCTGACTTACTA 827
Db 538 AAACCAATTTGCTATGATAGTGGCGCATGTCACTAACTTGAGAAAGCTGACTTACTA 597
QY 828 AAGGCTATTCAAGAAACAATTGATCGCTAAGTCCAGTAAACGACGACTCTTGAAGTGC 887
Db 598 AAGGCTATTCAAGAAACAATTGATCGCTAAGTCCAGTAAACGACGACTCTTGAAGTGC 657
QY 888 ATTCGATTTTTCGAAAGCGATGCAACCATTTCTGATCGAAACGCAAGTCTTCTTGCCTGAC 947
Db 658 ATTCGATTTTTCGAAAGCGATGCAACCATTTCTGATCGAAACGCAAGTCTTCTTGCCTGAC 717
QY 948 AAAGATGCTTCGGTAACTTTCGCGACCCGAACTGCTCAAGAAATTTTGTCTAAGCGGACAT 1007
Db 718 AAAGATGCTTCGGTAACTTTCGCGACCCGAACTGCTCAAGAAATTTTGTCTAAGCGGACAT 777
QY 1008 GTGCGGGTTAGACATTAAGAAACCAATTAAGAAACCAAGGAAATCTGTTGAAGTGC 1067
Db 778 GTGCGGGTTAGACATTAAGAAACCAATTAAGAAACCAAGGAAATCTGTTGAAGTGC 837
QY 1068 GAATATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTCAAGACAGTCTCAAA 1127
Db 838 GAATATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTCAAGACAGTCTCAAA 897
QY 1128 GATCTAAGCTATTGAAAAACATAGCTATCGGTGACACATCACTCTCAAGAAATTACTA 1187
Db 898 GATCTAAGCTATTGAAAAACATAGCTATCGGTGACACATCACTCTCAAGAAATTACTA 957
QY 1188 GCTGAAGCAACGATTTTAAACAAACACCCGCGCTATGATGATTTTGAACGCGAC 1247
Db 958 GCTGAAGCAACGATTTTAAACAAACACCCGCGCTATGATGATTTTGAACGCGAC 1017
QY 1248 TCCCTCAATCGTCACTGATGACATGACATTTCCGTAAGATTTTCAAGATCAAGAG 1307
Db 1018 TCCCTCAATCGTCACTGATGACATGACATTTCCGTAAGATTTTCAAGATCAAGAG 1077
QY 1308 TTTACTTACCGTGTAAAAATCGGGAACAAGCTTATAGATCAATAAAAAATCTGGTCTG 1367
Db 1078 TTTACTTACCGTGTAAAAATCGGGAACAAGCTTATAGATCAATAAAAAATCTGGTCTG 1137
QY 1368 AATGAAGAATAAACAACATGACCTGATCTCTGAGAAATTTACGTCCTTAAAAAAGGG 1427
Db 1138 AATGAAGAATAAACAACATGACCTGATCTCTGAGAAATTTACGTCCTTAAAAAAGGG 1197
QY 1428 GAAAAACCCGTATGCTCTTGTGATCGGAGTCACTTGAACCTGTTCAACATCAATTAAGTT 1487
Db 1198 GAAAAACCCGTATGCTCTTGTGATCGGAGTCACTTGAACCTGTTCAACATCAATTAAGTT 1257
QY 1488 GATGTCGATCAACAAGAAATGCTAAAAAGTGAAGAGCTCTTAAACGCTAAGGAAGCTTAC 1547
Db 1258 GATGTCGATCAACAAGAAATGCTAAAAAGTGAAGAGCTCTTAAACGCTAAGGAAGCTTAC 1317
QY 1548 TTAGACTTCAGAGATTTATTCGATCTCTGCTGATTAAGGCTTAACTCTTCAACAATCTC 1607
Db 1318 TTAGACTTCAGAGATTTATTCGATCTCTGCTGATTAAGGCTTAACTCTTCAACAATCTC 1377
QY 1608 GATGCTTTGGTATTTATGAGCTATACCTTAACGGAAGGAAGGATTAATCAAGATGAC 1667

```

```

Db 1378 GATGCTTTGGTATTAAGACTATACCTTAATCGAAAAAGTAGAGATTAATCAGATGAC 1437
QY 1668 ACCAAGCGTATCATTAACCGTTTATATGGCAAGGACCCGGAAGAGAGAAATGCTTAC 1727
Db 1438 ACCAAGCGTATCATTAACCGTTTATATGGCAAGGACCCGGAAGAGAGAAATGCTTAT 1497
QY 1728 CATTTAGCTGGTGGT 1742
Db 1498 CATTTAGCTTATGAT 1512

```

RESULT 13

AAAX16633
ID AAAX16633 standard; DNA; 1242 BP.

XX AAX16633;

DT 04-MAY-1999 (first entry)

XX Streptococcus equisimilis mutant streptokinase K59E encoding DNA.

XX Streptococcus equisimilis H46A; streptokinase; mutant; fibrinolytic;

XX plasma clot; hydrolysis; haemolytic Streptococcus; plasminogen; plasmin;

XX serine protease; fibrin; blood clot; thrombolytic;

XX vascular thromboembolytic symptom; acute myocardial infarction;

XX fibrinolysis; resistance; ds.

XX Streptococcus dysgalactiae subsp. equisimilis.

OS Synthetic.

XX Key

FT CDS

FT CDS

FT CDS

FT CDS

FT CDS

FT CDS

FT CDS

FT CDS

FT CDS

FT CDS

FT CDS

FT CDS

FT CDS

FT CDS

FT CDS

FT CDS

FT CDS

FT CDS

FT CDS

Query Match 54.4%; Score 1140.6; DB 2; Length 1242;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1146; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Sequence 1242 BP; 421 A; 268 C; 239 G; 314 T; 0 U; 0 Other;
The present invention describes a mutant streptokinase (SK) polypeptide in which at least one of the amino acids in the Pro58-Lys59-Ser60-Lys61 segment of the corresponding native SK is replaced by another amino acid. The present sequence encodes mutant SK K59E. SK is a secretory protein of haemolytic Streptococcus able to activate human plasminogen (Hpi) to plasmin (Hpin), which is a serine protease able to catalyse the hydrolysis of fibrin in blood clots. The SK is useful as a thrombolytic agent in the treatment of vascular thromboembolytic symptoms such as acute myocardial infarction. Compared with wild-type SK, the K59E mutant is more resistant to degradation by human plasmin and is more effective both in acting as a fibrinolytic agent and in activating human plasminogen.

```

QY 588 ATTGCTGAGACCTGAGTGGCTGCTGAGACCGTCCATCTGTCAACACAGCCCAATGGTGTG 647
DB 1 ATTGCTGAGACCTGAGTGGCTGCTGAGACCGTCCATCTGTCAACACAGCCCAATGGTGTG 60
QY 648 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGCATTAAGCTTAAATTTTGAATC 707
DB 61 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGCATTAAGCTTAAATTTTGAATC 120
QY 708 GATCTAACATCAGCAGCTGCTCATGAGAGAAAGACAGCAAGGCTTAAGTCCAAATCA 767
DB 121 GATCTAACATCAGCAGCTGCTCATGAGAGAAAGACAGCAAGGCTTAAGTCCGAATCA 180
QY 768 AAACCATTTGCTACTGATAGTGGCGCATGTCACTAACTTGAAGAAAGCTGACTTACTA 827
DB 181 AAACCATTTGCTACTGATAGTGGCGCATGTCACTAACTTGAAGAAAGCTGACTTACTA 240
QY 828 AAGGCTTATTCAGAGCAATTTGATGCTTAAGCTCCACAGTAAAGAGACTTCTTGAAGTC 887
DB 241 AAGGCTTATTCAGAGCAATTTGATGCTTAAGCTCCACAGTAAAGAGACTTCTTGAAGTC 300
QY 888 ATTGATTTTGCAGAGCAGTGCACCATTAAGTGAAGAGGGAAGGCTACTTTGCTGAC 947
DB 301 ATTGATTTTGCAGAGCAGTGCACCATTAAGTGAAGAGGGAAGGCTACTTTGCTGAC 360
QY 948 AAAGATGTTTCGGTAACTTTCGCCAGCCCAACCTGTCCAAGAAATTTTGTCTAAGCGAGAT 1007
DB 361 AAAGATGTTTCGGTAACTTTCGCCAGCCCAACCTGTCCAAGAAATTTTGTCTAAGCGAGAT 420
QY 1008 GTGCGGCTTGAACCATTTAAAGAAAAACAATACAAAACAGAGGAAATCTGTTGATGTG 1067
DB 421 GTGCGGCTTGAACCATTTAAAGAAAAACAATACAAAACAGAGGAAATCTGTTGATGTG 480
QY 1068 GAATATATCTGTACAGTTTACCTCTTAAACCTGATGAGATTTTCAAGCCAGGCTCTCAA 1127
DB 481 GAATATATCTGTACAGTTTACCTCTTAAACCTGATGAGATTTTCAAGCCAGGCTCTCAA 540
QY 1128 GATATCTAAGCTATTTGAAAAACATGACTATCGGTGACACCATCACTCAAGAAATTTACTA 1187
DB 541 GATATCTAAGCTATTTGAAAAACATGACTATCGGTGACACCATCACTCAAGAAATTTACTA 600
QY 1188 GCTTAAAGCAGAAAGCATTTTAAACAAAACCAAGGCTATAGCAATTTTAAAGCGTGAC 1247
DB 601 GCTTAAAGCAGAAAGCATTTTAAACAAAACCAAGGCTATAGCAATTTTAAAGCGTGAC 660
QY 1248 TCCCTCAATCGTCACTGATGACATGACATTTCCGTACGATTTTAAACATGATCAAGAG 1307
DB 661 TCCCTCAATCGTCACTGATGACATGACATTTCCGTACGATTTTAAACATGATCAAGAG 720
QY 1308 TTTACTTAAACCGTGTAAATAATCGGGAACAAGCTTATAGATCAATTAATAAATCTGCTG 1367
DB 721 TTTACTTAAACCGTGTAAATAATCGGGAACAAGCTTATAGATCAATTAATAAATCTGCTG 780
QY 1368 AATGAAGAAATAAACAACACTGACCTGATCTGTGAGAAATTTTACGTCCTTAAAAAGGG 1427
DB 781 AATGAAGAAATAAACAACACTGACCTGATCTGTGAGAAATTTTACGTCCTTAAAAAGGG 840
QY 1428 GAAAAACCGTATGATCCCTTTGATCGAGTCACTTGAATCTGTTACCAATCAAAATACGTT 1487
DB 841 GAAAAACCGTATGATCCCTTTGATCGAGTCACTTGAATCTGTTACCAATCAAAATACGTT 900
QY 1488 GATGTGATACCAACGAATTTGCTAAAGAGAGAGAGCTTTTAAACAGCTAGGAAAGCTAAC 1547
DB 901 GATGTGATACCAACGAATTTGCTAAAGAGAGAGAGCTTTTAAACAGCTAGGAAAGCTAAC 960
QY 1548 TTAGACTTCAGAGATTTTATAGATCTCGTGAATAGGCTTAACTACTTACACAAATCTC 1607
DB 961 TTAGACTTCAGAGATTTTATAGATCTCGTGAATAGGCTTAACTACTTACACAAATCTC 1020
QY 1608 GATGCTTTTGGTATTTATGACATTAACCTTAACTGGAAGAGAGATTAATCAAGATGAC 1667
DB 1021 GATGCTTTTGGTATTTATGACATTAACCTTAACTGGAAGAGAGATTAATCAAGATGAC 1080

```

```

QY 1668 ACCAACCGTATCATTAACCGTTTATATGAGGACAGACCCGAGAGAGAGATGCTACTAC 1727
DB 1081 ACCAACCGTATCATTAACCGTTTATATGAGGACAGACCCGAGAGAGAGATGCTACTAC 1140
QY 1728 CATTTAGCTGTGTGT 1742
DB 1141 CATTTAGCTGTGTAT 1155

```

```

RESULT 14
AAA37628
AAA37628 standard; DNA; 1327 BP.

```

```

XX AAA37628;
AC 15-SEP-2003 (revised)
DT 13-OCT-2000 (first entry)
DE Streptokinase-NTR gene.

```

```

KW Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
KW plasminogen; human; fibrinectin; thrombolytic therapy;
KW cardiovascular disorder; ss.

```

```

XX Streptococcus dysgalactiae subsp. equisimilis.

```

```

PN EP1024192-A2.

```

```

PD 02-AUG-2000.

```

```

XX 23-DEC-1999; 99EP-00310541.

```

```

PR 24-DEC-1998; 98IN-DE003825.

```

```

XX (COUL ) CSIR COUNCIL SCI IND RES.

```

```

PI Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
PI Yadav M;

```

```

XX WPI; 2000-516032/47.

```

```

PT Hybrid streptokinase-fibrin binding domain polypeptides useful for
PT thrombolytic therapy comprises a streptokinase fused with fibrin binding
PT domains of human fibrinectin.

```

```

XX Example 1; Fig 14; 58bp; English.

```

```

XX This sequence represents a streptokinase-NTR (SK-NTR) gene (where NTR
XX stands for N-terminally repaired with native sequence). The invention
XX relates to a hybrid plasminogen activator (PA) comprises a polypeptide
XX fusion between streptokinase (SK), which are capable of plasminogen (PG)
XX activation, and fibrin binding regions of human fibrinectin, which are
XX from fibrin binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA
XX possesses the ability to bind with fibrin independently and also
XX characteristically retains a PG activation ability which becomes evident
XX only after a pronounced duration, or lag, after exposure of the PA to a
XX suitable animal or human PG. The hybrid streptokinase-fibrin binding
XX domain polypeptides are useful in thrombolytic therapy for various kinds
XX of cardiovascular disorders. The hybrids have enhanced fibrin selectivity
XX as well as kinetics of plasminogen activation that are distinct from that
XX of natural streptokinase in being characterised by a temporary delay, or
XX lag of several minutes in the natural rate of the catalytic conversion of
XX plasminogen to plasmin (i.e. delayed-action thrombolytic). The proteins
XX can bind tightly with fibrin in blood clots soon after introduction into
XX the vascular system without significantly activating the circulating
XX blood plasminogen to plasmin, thus aiding in the localisation of the
XX plasminogen activation process to the site of pathological thrombus. This
XX overcomes systemic plasminogen activation encountered during clinical use
XX of streptokinase. (Updated on 15-SEP-2003 to standardise OS field)

```

```

SQ Sequence 1327 BP; 456 A; 279 C; 250 G; 342 T; 0 U; 0 Other;

```

```

Query Match 54.1%; Score 1134.2; DB 3; Length 1327;

```

Best Local Similarity 98.9%; Pred. No. 2.3e-313;
Matches 1142; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

```

QY 588 ATTGCTGAGACCTGATGCGCTGCTAGACCTGCTCACTCTGTCAACACAGCGCAATGGTGT 647
DB 83 ATAGCTGCTGCTGATATGCTAGTATGATGCTCTTCTTAATAATACAGCAATGGTGT 142
QY 648 AGGTTGCTGCTGATCTGTGAGGGGACGAATCAAGACATTAAGTCTTAATTTTGAATC 707
DB 143 AGGTTGCTGCTGATCTGTGAGGGGACGAATCAAGACATTAAGTCTTAATTTTGAATC 202
QY 708 GATCTTAACATCAGACCTGCTCATGAGAAAGACAGCAAGGCTTAAGTCAAAAATCA 767
DB 203 GATCTTAACATCAGACCTGCTCATGAGAAAGACAGCAAGGCTTAAGTCAAAAATCA 262
QY 768 AAACCATTTGCTACTGATAGTGGCGCATGTCACTAACTTGAGAAAGTGACTTACTA 827
DB 263 AAACCATTTGCTACTGATAGTGGCGCATGTCACTAACTTGAGAAAGTGACTTACTA 322
QY 828 AAGGCTATTCAAGAACTTGAATGCTAAGTCCAGTAAAGCACTACTTGAAGTC 887
DB 323 AAGGCTATTCAAGAACTTGAATGCTAAGTCCAGTAAAGCACTACTTGAAGTC 382
QY 888 ATTGATTTTGCAGAGCGATGCAACCATTAATGATGAAACGGCAAGTCTACTTTGCTGAC 947
DB 383 ATTGATTTTGCAGAGCGATGCAACCATTAATGATGAAACGGCAAGTCTACTTTGCTGAC 442
QY 948 AAAGATGCTTCCGTTAACTTCCGACCCAACTGTCTCAAGAAATTTTGTAAAGCGGACAT 1007
DB 443 AAAGATGCTTCCGTTAACTTCCGACCCAACTGTCTCAAGAAATTTTGTAAAGCGGACAT 502
QY 1008 GTGGCGGTTAGACCATTAATAAGAAACCAATACAAACCAAGGAATCTGTGATGTG 1067
DB 503 GTGGCGGTTAGACCATTAATAAGAAACCAATACAAACCAAGGAATCTGTGATGTG 562
QY 1068 GAATATCTGTATACAGTTTACTCTCTTAAACCTGTATGACGATTTTCAAGCGAGTCTCAA 1127
DB 563 GAATATCTGTATACAGTTTACTCTCTTAAACCTGTATGACGATTTTCAAGCGAGTCTCAA 622
QY 1128 GATCTAAGCTATTTGAAACACATAGCTATCGGTGACACATCACTCAAGAAATTA 1187
DB 623 GATCTAAGCTATTTGAAACACATAGCTATCGGTGACACATCACTCAAGAAATTA 682
QY 1188 GCTCAAGCACAAAGCATTTTAAACAAACCAACCGCTATACGATTTTGAACGTCGAC 1247
DB 683 GCTCAAGCACAAAGCATTTTAAACAAACCAACCGCTATACGATTTTGAACGTCGAC 742
QY 1248 TCCCTCAATGCTCATCATGACATGACATTTTCCGTACGATTTTCAACAAATGATCAAGAG 1307
DB 743 TCCCTCAATGCTCATCATGACATGACATTTTCCGTACGATTTTCAACAAATGATCAAGAG 802
QY 1308 TTTACTTACCGTGTAAATAATCGGGAAACAAGCTTATAGATCATATAAAATCTGTCTG 1367
DB 803 TTTACTTACCGTGTAAATAATCGGGAAACAAGCTTATAGATCATATAAAATCTGTCTG 862
QY 1368 AATGAAGAATTAACAACACTGACCTGATCTTGAAGAAATTAATGCTTTAAAAAGGG 1427
DB 863 AATGAAGAATTAACAACACTGACCTGATCTTGAAGAAATTAATGCTTTAAAAAGGG 922
QY 1428 GAAAACCCGTAATGCTTGTATGCGAGTCACTTGAACCTGTTCACATCAATAATGAGTT 1487
DB 923 GAAAACCCGTAATGCTTGTATGCGAGTCACTTGAACCTGTTCACATCAATAATGAGTT 982
QY 1488 GATGTCGATACCAAGCAATGCTTAAAGTGAAGCACTTAAACAGCTAAGCAAGTAAAC 1547
DB 983 GATGTCGATACCAAGCAATGCTTAAAGTGAAGCACTTAAACAGCTAAGCAAGTAAAC 1042
QY 1548 TTAGACTTCAAGATTTATACGATCTCTGTGATTAAGGCTTAACTACTTCAACAAATCTC 1607
DB 1043 TTAGACTTCAAGATTTATACGATCTCTGTGATTAAGGCTTAACTACTTCAACAAATCTC 1102
QY 1608 GATGCTTTGTATTAATGACTATACCTTAACTGAAAGATGAAGATTAATCAAGATGAC 1667

```

```

DB 1103 GATGCTTTGGTATATGACTATACCTTAACCTGGAAGAAAGTAGAGATTAATCAAGATGAC 1162
QY 1668 ACCAACCTTATCATTAACCGTTTATATGCGCAAGCAACCCGAAGAGAAATGCTATAC 1727
DB 1163 ACCAACCTTATCATTAACCGTTTATATGCGCAAGCAACCCGAAGAGAAATGCTATAT 1222
QY 1728 CATTTAGCTGGTGT 1742
DB 1223 CATTTAGCTATGAT 1237

```

RESULT 15

AAT77778
ID AAT77778 standard; cDNA; 2566 BP.

XX AAT77778;

AC 17-OCT-2003 (revised)

DT 01-OCT-1997 (first entry)

XX Coding sequence for plasminogen-binding fragment of Streptokinase.

XX Plasminogen-binding fragment; streptokinase; degradation; MBP;

XX thrombolytic agent; blood clot; bolus; maltose-binding protein; ds.

XX Streptococcus dysgalactiae subsp. equisimilis.

XX W09641883-A1.

XX 27-DEC-1996.

XX 07-JUN-1996; 96WO-US009640.

XX 09-JUN-1995; 95US-00488940.

XX (HARD) HARVARD COLLEGE.

XX Read GLI;

XX WPI; 1997-065469/06.

XX Modified forms of streptokinase resistant to enzymatic cleavage - useful

XX as thrombolytic agents in treating thrombosis and in medical equipment.

XX Example 1; Page 22-23; 65pp; English.

XX This sequence encodes the wild type plasminogen-binding fragment of
CC streptokinase. The protein fragment encoded by this sequence was used in
CC the design of a modified streptokinase has an in vitro degradation rate
CC at least 2 times slower than that of native streptokinase. Compounds
CC containing modified streptokinases are specifically used as thrombolytic
CC agents for dissolving blood clots in vivo in a mammal, preferably at a
CC dose of 20000 U/kg, opt. as a bolus rather than by continuous infusion.
XX (Updated on 17-OCT-2003 to standardise OS field)

XX Sequence 2566 BP; 825 A; 545 C; 469 G; 727 T; 0 U; 0 Other;

XX Query Match 54.0%; Score 1132.4; DB 2; Length 2566;

XX Best Local Similarity 99.3%; Pred. No. 1.1e-312;

XX Matches 1148; Conservative 0; Mismatches 6; Indels 2; Gaps 1;

QY 587 TATTGCTGACCTGATGAGTGTGCTAGACCGTTCATCTGTCAACAAACGCAATGGTGT 646

DB 896 TATTGCTGACCTGATGAGTGTGCTAGACCGTTCATCTGTCAACAAACGCAATGAGTGT 955

QY 647 TAGGTTGCTGCTACTGTGAGGGGACGAATCAAGACATTAAGTCTTAATTTTGAAT 706

DB 956 TAGGTTGCTGCTACTGTGAGGGGACGAATCAAGACATTAAGTCTTAATTTTGAAT 1015

QY 707 CGATCTTAACATCAGACCTGCTCATGAGAAAGACAGCAAGGCTTAAGTCAAAAATC 766

DB 1016 CGATCTTAACATCAGACCTGCTCAT--AGAAAGACAGCAAGGCTTAAGTCAAAAATC 1073

QY 767 AAAACATTGCTACTGATAGTGGCGGAGTGTCAATAACTTGAAAGCTGACTTACT 826
| | | | |
Db 1074 AAAACATTGCTACTGATAGTGGCGGAGTGTCAATAACTTGAAAGCTGACTTACT 1133
| | | | |
QY 827 AAAGGCTATTCAGAAACAATTGATCGCTAAAGTCCAGTAAAGCACTACTTTGAGGT 886
| | | | |
Db 1134 AAAGGCTATTCAGAAACAATTGATCGCTAAAGTCCAGTAAAGCACTACTTTGAGGT 1193
| | | | |
QY 887 CATGTATTTGCAAGCAGTGCACCAATTATCTGATCGAAACGCAAGTCTACTTTGCTGA 946
| | | | |
Db 1194 CATGTATTTGCAAGCAGTGCACCAATTATCTGATCGAAACGCAAGTCTACTTTGCTGA 1253
| | | | |
QY 947 CAAAGATGTTCCGTTAACTTCCGACCCAACTGTCCAGATTTTGTAAAGCGACA 1006
| | | | |
Db 1254 CAAAGATGTTCCGTTAACTTCCGACCCAACTGTCCAGATTTTGTAAAGCGACA 1313
| | | | |
QY 1007 TGTCCCGGTAGACCATATATAAGAAAAACAATACAAAACCAAGGAAATCTTTGATGT 1066
| | | | |
Db 1314 TGTCCCGGTAGACCATATATAAGAAAAACAATACAAAACCAAGGAAATCTTTGATGT 1373
| | | | |
QY 1067 GGAATATACTGTACAGTTTACTCCCTTAAACCTGATGAGATTTGAGCAAGGCTCAA 1126
| | | | |
Db 1374 GGAATATACTGTACAGTTTACTCCCTTAAACCTGATGAGATTTGAGCAAGGCTCAA 1433
| | | | |
QY 1127 AGATTAAGCTATTTGAAAACTAGCTATCGTGCACACATCAATCTCAAGAAATTACT 1186
| | | | |
Db 1434 AGATTAAGCTATTTGAAAACTAGCTATCGTGCACACATCAATCTCAAGAAATTACT 1493
| | | | |
QY 1187 AGCTCAAGCAAAAGCAATTTTAAACAAAACCAACCGGCTATACGATTTATGAACTGA 1246
| | | | |
Db 1494 AGCTCAAGCAAAAGCAATTTTAAACAAAACCAACCGGCTATACGATTTATGAACTGA 1553
| | | | |
QY 1247 CTGCTCAATGCTGACATGACATGACATTTTCGTAAGATTTTAAACGATGACAGA 1306
| | | | |
Db 1554 CTGCTCAATGCTGACATGACATGACATTTTCGTAAGATTTTAAACGATGACAGA 1613
| | | | |
QY 1307 GTTTACTTACCGTGTAAAAATCGGAAACAAGTTATAGATCAATAAAAAATCTGCTCT 1366
| | | | |
Db 1614 GTTTACTTACCGTGTAAAAATCGGAAACAAGTTATAGATCAATAAAAAATCTGCTCT 1673
| | | | |
QY 1367 GAAATGAAGAAATTAACAACACTGACCTGATCTGAGAAATATTACGCTTAAAAAAG 1426
| | | | |
Db 1674 GAAATGAAGAAATTAACAACACTGACCTGATCTGAGAAATATTACGCTTAAAAAAG 1733
| | | | |
QY 1427 GGAAGAGCCGTATGATCCCTTTGATCGGACGTCATTGAAACTGTTCAACATCAATACGT 1486
| | | | |
Db 1734 GGAAGAGCCGTATGATCCCTTTGATCGGACGTCATTGAAACTGTTCAACATCAATACGT 1793
| | | | |
QY 1487 TGATGTGATACCAACGAATTTGCTAAAAAGTGAAGCAGCTTAAACGCTAGCGAACGTA 1546
| | | | |
Db 1794 TGATGTGATACCAACGAATTTGCTAAAAAGTGAAGCAGCTTAAACGCTAGCGAACGTA 1853
| | | | |
QY 1547 CTTAGACTTCAGAGATTTATACGATCCTCGTATAGGCTAAACTACTCTAACAAATCT 1606
| | | | |
Db 1854 CTTAGACTTCAGAGATTTATACGATCCTCGTATAGGCTAAACTACTCTAACAAATCT 1913
| | | | |
QY 1607 CGATGCTTTTGTATTATGACATATACCTTAACTGAAAAAGTAGAGATAATCAAGATGA 1666
| | | | |
Db 1914 CGATGCTTTTGTATTATGACATATACCTTAACTGAAAAAGTAGAGATAATCAAGATGA 1973
| | | | |
QY 1667 CACCAACCGTATCATTAACCGTTTATATGCGAAGCAGCCGAAAGAGAAATGCTAGCTA 1726
| | | | |
Db 1974 CACCAACCGTATCATTAACCGTTTATATGCGAAGCAGCCGAAAGAGAAATGCTAGCTA 2033
| | | | |
QY 1727 CCATTTAGCTGTGTGT 1742
| | | | |
Db 2034 TCATTTAGCCTATGAT 2049
| | | | |

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2006, 16:24:19 ; Search time 370.056 Seconds
(without alignments)
10068.109 Million cell updates/sec

Title: US-09-940-235-12

Perfect score: 2096
Sequence: 1 cgaagaccatcatgttgcgtt.....cctcaccgattcgttag 2096

Scoring table:

IDENTITY NUC
Gapop 10-0, Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 10%

Listing first 45 summaries

Database :

Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/1.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP.COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RG.COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1150.8	54.9	2385	3	US-09-211-542A-1
2	1145.4	54.6	1242	3	US-09-211-542A-5
3	1143.8	54.6	1242	2	US-08-568-393B-1
4	1140.6	54.4	1242	2	US-08-568-393B-2
5	1132.4	54.0	2562	2	US-08-488-940-19
6	1125.4	53.7	1458	2	US-07-854-596B-42
7	1124.6	53.7	1245	2	US-07-703-778D-1
8	1122.4	53.5	1335	2	US-07-854-596B-14
9	1122.4	53.5	1512	2	US-07-854-596B-27
10	1122.2	53.5	2589	2	US-07-854-596B-34
11	1121.4	53.5	1257	2	US-07-854-596B-25
12	1121.4	53.5	1317	2	US-07-854-596B-18
13	1121.4	53.5	1467	2	US-07-854-596B-46
14	1093.6	52.2	1122	3	US-09-374-038-13
15	1093.6	52.2	1122	3	US-09-658-179-13
16	1093.6	52.2	1158	3	US-09-374-038-14
17	1093.6	52.2	1158	3	US-09-658-179-14
18	1086	51.8	1209	3	US-09-374-038-11
19	1086	51.8	1209	3	US-09-658-179-11
20	1086	51.8	1245	3	US-09-374-038-12
21	1086	51.8	1245	3	US-09-658-179-12
22	1077.2	51.4	2253	2	US-07-854-596B-39
23	1076	51.3	1119	2	US-07-854-596B-30
24	973	46.4	1068	3	US-09-211-542A-11

25	970	46.3	2208	3	US-09-211-542A-3	Sequence 3, Appl
26	759.4	36.2	1242	9	5240845-2	Patent No. 5240845
27	759.4	36.2	1262	9	5240845-3	Patent No. 5240845
28	450	21.5	450	3	US-09-211-542A-13	Sequence 13, Appl
29	330	15.7	7679	3	US-09-220-132-38	Sequence 38, Appl
30	330	15.7	7680	3	US-09-023-655-1289	Sequence 1289, Ap
31	330	15.7	7680	6	PCT-US95-09819-6	Sequence 6, Appl
32	330	15.7	7803	6	PCT-US93-12687-1	Sequence 1, Appl
33	330	15.7	8044	3	US-09-566-921-135	Sequence 135, App
34	330	15.6	7705	2	US-08-259-569-16	Sequence 16, Appl
35	326.8	15.6	7705	2	US-08-826-885-16	Sequence 16, Appl
36	326.8	15.6	7705	9	545158-2	Patent No. 545158
37	326.8	15.6	7705	9	545158-2	Sequence 13, Appl
38	279.8	13.3	4811	3	US-09-813-718-13	Sequence 15, Appl
39	278.8	13.3	4742	3	US-09-813-718-15	Sequence 11, Appl
40	278.4	13.3	4877	3	US-09-813-718-11	Sequence 9, Appl
41	278.4	13.3	5018	3	US-09-813-718-9	Sequence 5, Appl
42	278	13.3	4100	3	US-09-813-718-5	Sequence 3, Appl
43	278	13.3	4682	3	US-09-813-718-7	Sequence 7, Appl
44	278	13.3	4682	3	US-09-813-718-7	Sequence 1, Appl
45	278	13.3	5174	3	US-09-813-718-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-09-211-542A-1
Sequence 1, Application US/09211542A
Patent No. 6210667
GENERAL INFORMATION:
APPLICANT: Reed, Guy L.
TITLE OF INVENTION: BACTERIAL FIBRIN-DEPENDENT PLASMINOGEN ACTIVATOR
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESS: BROMBERG & SUNSTEIN, LLP
STREET: 125 Summer Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/211,542A
FILING DATE: 15-December-1998
CLASSIFICATION: 1653
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/069,497
FILING DATE: 15-December-1997
ATTORNEY/AGENT INFORMATION:
NAME: Attorney, Strimpel, Harriet M.
REGISTRATION NUMBER: 37,008
REFERENCE/DOCKET NUMBER: 1874/111
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 443-9292
TELEFAX: (617) 443-0064
INFORMATION FOR SEQ ID NO. 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2385 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2385
US-09-211-542A-1
Query Match 54.9%, Score 1150.8, DB 3, Length 2385,

Best Local Similarity 99.0%; Pred. No. 0;
Matches 1158; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

```
Oy 573 TTACCGATGTTGATTTGCTGACCTGAGTGGCTGTAACCGTCCATCTGTCAACAC 632
    |||
Db 1129 TCCATGAGGGTGGATTTGCTGACCTGAGTGGCTGTAACCGTCCATCTGTCAACAC 1188
    |||
Oy 633 AGCCAAATGGTGTAGCGTTGCTGTAAGTGGGGAAGCAATCAAGCATTAAGTCTT 692
    |||
Db 1189 AGCCAAATGATGTTAGCGTTGCTGTAAGTGGGGAAGCAATCAAGCATTAAGTCTT 1248
    |||
Oy 693 AAATTTTGAATTCGATCTTAACATCAACGACCTGCTCATGAGGGAAGACAGCAAGGC 752
    |||
Db 1249 AAATTTTGAATTCGATCTTAACATCAACGACCTGCTCATGAGGGAAGACAGCAAGGC 1308
    |||
Oy 753 TTAAGTCCAAATCAAAACCATTTGCTAAGTATGTCGCGGATGTCACTAACTTGAG 812
    |||
Db 1309 TTAAGTCCAAATCAAAACCATTTGCTAAGTATGTCGCGGATGTCACTAACTTGAG 1368
    |||
Oy 813 AAAGCTGACTTACTTAAGGCTATCAAGAACATTTGATGCTAACGTCACAGTAACGAC 872
    |||
Db 1369 AAAGCTGACTTACTTAAGGCTATCAAGAACATTTGATGCTAACGTCACAGTAACGAC 1428
    |||
Oy 873 GACTACTTTGAGGTCATTTGATTTTTCGAGGCAACCATTAAGTCAAGGCAAG 932
    |||
Db 1429 GACTACTTTGAGGTCATTTGATTTTTCGAGGCAACCATTAAGTCAAGGCAAG 1488
    |||
Oy 933 GTCTACTTTGCTGACAAAGATGTTGCGTAACTTTCGACCAACCTGTCCAAAGATT 992
    |||
Db 1489 GTCTACTTTGCTGACAAAGATGTTGCGTAACTTTCGACCAACCTGTCCAAAGATT 1548
    |||
Oy 993 TTGCTAAGCGGACATGTCGCGTAAACCATTAAGGGAAGCAATCAAGGCAAGC 1052
    |||
Db 1549 TTGCTAAGCGGACATGTCGCGTAAACCATTAAGGGAAGCAATCAAGGCAAGC 1608
    |||
Oy 1053 AAATCTGTTGATGTAATTAAGTCTGTAAGTTCCTCCCTTAACCGTGAAGCATTTTC 1112
    |||
Db 1609 AAATCTGTTGATGTAATTAAGTCTGTAAGTTCCTCCCTTAACCGTGAAGCATTTTC 1668
    |||
Oy 1113 AGACGAGTCTCAAGATTAAGTTCGTAAGTTCGTAAGTTCGTAAGTTCGTAAGTTC 1172
    |||
Db 1669 AGACGAGTCTCAAGATTAAGTTCGTAAGTTCGTAAGTTCGTAAGTTCGTAAGTTC 1728
    |||
Oy 1173 TCTCAAGATTAAGTTCGTAAGTTCGTAAGTTCGTAAGTTCGTAAGTTCGTAAGTTC 1232
    |||
Db 1729 TCTCAAGATTAAGTTCGTAAGTTCGTAAGTTCGTAAGTTCGTAAGTTCGTAAGTTC 1788
    |||
Oy 1233 ATTATGAAGTGAATCTGTAAGTTCGTAAGTTCGTAAGTTCGTAAGTTCGTAAGTTC 1292
    |||
Db 1789 ATTATGAAGTGAATCTGTAAGTTCGTAAGTTCGTAAGTTCGTAAGTTCGTAAGTTC 1848
    |||
Oy 1293 CCAATGATCAAGATTAAGTTCGTAAGTTCGTAAGTTCGTAAGTTCGTAAGTTCGTAAGTTC 1352
    |||
Db 1849 CCAATGATCAAGATTAAGTTCGTAAGTTCGTAAGTTCGTAAGTTCGTAAGTTCGTAAGTTC 1908
    |||
Oy 1353 AAAAAATCTGCTGTAAGTTCGTAAGTTCGTAAGTTCGTAAGTTCGTAAGTTCGTAAGTTC 1412
    |||
Db 1909 AAAAAATCTGCTGTAAGTTCGTAAGTTCGTAAGTTCGTAAGTTCGTAAGTTCGTAAGTTC 1968
    |||
Oy 1413 GTCTTAAAAAAGGGGAAAAAGCGGTATGATCCCTTTGATGCACTGCACTGTAAGTTC 1472
    |||
Db 1969 GTCTTAAAAAAGGGGAAAAAGCGGTATGATCCCTTTGATGCACTGCACTGTAAGTTC 2028
    |||
Oy 1473 ACCATCAATATCGTGAATGTCATCAACGAAATTTGTAAGTTCGTAAGTTCGTAAGTTC 1532
    |||
Db 2029 ACCATCAATATCGTGAATGTCATCAACGAAATTTGTAAGTTCGTAAGTTCGTAAGTTC 2088
    |||
Oy 1533 GCTAGCGAAGCTTAAGTTCGTAAGTTCGTAAGTTCGTAAGTTCGTAAGTTCGTAAGTTC 1592
    |||
Db 2089 GCTAGCGAAGCTTAAGTTCGTAAGTTCGTAAGTTCGTAAGTTCGTAAGTTCGTAAGTTC 2148
    |||
Oy 1593 CTCTACACATCTCGATGCTTTGTTATGATCTTAAGTTCGTAAGTTCGTAAGTTCGTAAGTTC 1652
    |||
```

```
Db 2149 CTCTACACATCTCGATGCTTTGTTATGATCTTAAGTTCGTAAGTTCGTAAGTTCGTAAGTTC 2208
    |||
Oy 1653 GATATACGATGACACCAACCGTATCAACCGTTCATGCGCAACCGCAAGGA 1712
    |||
Db 2209 GATATACGATGACACCAACCGTATCAACCGTTCATGCGCAACCGCAAGGA 2268
    |||
Oy 1713 GAGAATGCTAGCTACCATTTAGCTGTGTT 1742
    |||
Db 2269 GAGAATGCTAGCTACCATTTAGCTGTGTT 2298
    |||
```

RESULT 2

US-09-211-542A-5
; Sequence 5, Application US/09211542A

; Patent No. 6210667

; GENERAL INFORMATION:

; APPLICANT: Reed, Guy L.

; TITLE OF INVENTION: BACTERIAL FIBRIN-DEPENDENT PLASMINOGEN ACTIVATOR

; NUMBER OF SEQUENCES: 14

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: BROMBERG & SUNSTEIN, LLP

; STREET: 125 Summer Street

; CITY: Boston

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02110

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/211,542A

; FILING DATE: 15-December-1998

; CLASSIFICATION: 1653

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/069,497

; FILING DATE: 15-December-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Attorney, Strimpel, Harriet M.

; REGISTRATION NUMBER: 37,008

; REFERENCE/DOCKET NUMBER: 1874/111

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617)443-9292

; TELEFAX: (617)443-0004

; INFORMATION FOR SEQ ID NO: 5:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1242 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 1..1242

; US-09-211-542A-5

Query Match 54.6%; Score 1145.4; DB 3; Length 1242;

Best Local Similarity 99.5%; Pred. No. 0;

Matches 1149; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Oy 588 ATTGCTGAGCTGAGTGGCTGCTGAGACCGTCCATCTGTCAACACAGCCCAATTTGTTT 647
    |||
Db 1 ATTGCTGAGCTGAGTGGCTGCTGAGACCGTCCATCTGTGTCAACACAGCCCAATTTGTTT 60
    |||
Oy 648 AGCGTTGCTGATCTGTTGAAGGGAAGCAATCAAGCATTAAGTTCATTAATTTTGAATTC 707
    |||
Db 61 AGCGTTGCTGATCTGTTGAAGGGAAGCAATCAAGCATTAAGTTCATTAATTTTGAATTC 120
    |||
Oy 708 GATCTAACATACGACCTGCTCATGAGGAAAGACAGACGCAAGCTTAAGTTCATAATCA 767
    |||
Db 121 GATCTAACATACGACCTGCTCATGAGGAAAGACAGACGCAAGCTTAAGTTCATAATCA 180
    |||
```

QY 768 AAACATTGCTAGTAGTGGCGGATGTCATTAACCTTGAGAAAGCTGACTACTA 827
DB 181 AAACATTGCTAGTAGTGGCGGATGTCATTAACCTTGAGAAAGCTGACTACTA 240
QY 828 AAGGCTATTCAGAAACAAATTGATGGCTTAAGTCCCACTTAACGAGACTACTTTAGAGTC 887
DB 241 AAGGCTATTCAGAAACAAATTGATGGCTTAAGTCCCACTTAACGAGACTACTTTAGAGTC 300
QY 888 ATTGATTTTTCAGAGCAGTGCATTAAGTATGATGCAAAAGGCAAGGCTGACTTTGCTGAC 947
DB 301 ATTGATTTTTCAGAGCAGTGCATTAAGTATGATGCAAAAGGCAAGGCTGACTTTGCTGAC 360
QY 948 AAAGATGCTTCGCTAACCTTGGCGAACCCCACTGTCAGAAATTTTCTAGACGAGAT 1007
DB 361 AAAGATGCTTCGCTAACCTTGGCGAACCCCACTGTCAGAAATTTTCTAGACGAGAT 420
QY 1008 GTGCGGCTTGAAGCAATTAAGAAACCAATACAAACCAAGGCAAAATCTGTTGATGTG 1067
DB 421 GTGCGGCTTGAAGCAATTAAGAAACCAATACAAACCAAGGCAAAATCTGTTGATGTG 480
QY 1068 GAATATATCTGTAAGTAACTTCCCTTAAGCCCTGATGATGCAATTCAGACGAGTCTCAA 1127
DB 481 GAATATATCTGTAAGTAACTTCCCTTAAGCCCTGATGATGCAATTCAGACGAGTCTCAA 540
QY 1128 GATATTAAGCTATTTGAAACCACTAGCTATCGGTGACACCATCATCTCAAGAAATTACTA 1187
DB 541 GATATTAAGCTATTTGAAACCACTAGCTATCGGTGACACCATCATCTCAAGAAATTACTA 600
QY 1188 GCTCAAGCACAAGCAATTTTAAACAAACCAACCCAGGCTATGAGATTTATGAGAGTGAC 1247
DB 601 GCTCAAGCACAAGCAATTTTAAACAAACCAACCCAGGCTATGAGATTTATGAGAGTGAC 660
QY 1248 TCCTCAATGCTCACTGATGCAATGATGATTTCCGTGATGATTTTCCAGAAAGATGATGAC 1307
DB 661 TCCTCAATGCTCACTGATGCAATGATGATTTCCGTGATGATTTTCCAGAAAGATGATGAC 720
QY 1308 TTTACTTACCGGTGTTAAATTCGGGAAACAAGCTTATGAGATCAATTAATAAATCTGCTG 1367
DB 721 TTTACTTACCGGTGTTAAATTCGGGAAACAAGCTTATGAGATCAATTAATAAATCTGCTG 780
QY 1368 AATGAAGAAATTAACCACTGATGCTGATGATTTTCCGTGATGATTTTCCAGAAAGATGATGAC 1427
DB 781 AATGAAGAAATTAACCACTGATGCTGATGATTTTCCGTGATGATTTTCCAGAAAGATGATGAC 840
QY 1428 GAAAACCCGATGATGCTTGTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 1487
DB 841 GAAAACCCGATGATGCTTGTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 900
QY 1488 GATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 1547
DB 901 GATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 960
QY 1548 TTAGACTTGAAGATTTTATGATGCTGCTGATGATGATGATGATGATGATGATGATGATGCT 1607
DB 961 TTAGACTTGAAGATTTTATGATGCTGCTGATGATGATGATGATGATGATGATGATGATGCT 1020
QY 1608 GATGCTTTTGTGATTTATGATGCTTAACTTGAAGAAAGTAAAGATATGATGATGATGATG 1667
DB 1021 GATGCTTTTGTGATTTATGATGCTTAACTTGAAGAAAGTAAAGATATGATGATGATGATG 1080
QY 1668 ACCAACCCTATCAATACCGTTTATATGAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCA 1727
DB 1081 ACCAACCCTATCAATACCGTTTATATGAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCA 1140
QY 1728 CATTAGCTGCTGCT 1742
DB 1141 CATTAGCTGCTGCT 1155

RESULT 3
US-08-568-393B-1
; Sequence 1, Application US/08568393B
; Patent No. 5876999

GENERAL INFORMATION:
APPLICANT: Hua-Lin Wu
APPLICANT: Guey-Yueh Shi
TITLE OF INVENTION: Preparation of novel streptokinase
TITLE OF INVENTION: mutants as improved thrombolytic agents
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSER: Jeling & Chang
STREET: Two No. 5876999th Second Street, Suite 290
CITY: San Jose
STATE: California
COUNTRY: USA
ZIP: 95113
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44MB
MEDIUM TYPE: storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect 6.1 on Window 3.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/568,393B
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Chi-Ping Chang
REGISTRATION NUMBER: 37,798
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (408) 288-8585
TELEFAX: (408) 288-8386
INFORMATION FOR SEO ID NO. 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1242 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE:
HYPOTHEICAL: N
ANTI-SENSE: N
ORIGINAL SOURCE:
ORGANISM: Streptococcus equisimilis H46A
INDIVIDUAL ISOLATE: Malke, H., Roe, B., and Ferretti, J. J.
INDIVIDUAL ISOLATE: "Nucleotide sequence of the streptokinase gene from Streptococcus equisimilis H46A"
CELL TYPE: Streptococcus equisimilis H46A
US-08-568-393B-1
Query Match 54.6%; Score 1143.8; DB 2; Length 1242;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1148; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 588 ATTGCTGACCTGAGTGGCTGCTGACCGTCCATCTGTCACAAAGCAATTTGTTGTT 647
DB 1 ATTGCTGACCTGAGTGGCTGCTGACCGTCCATCTGTCACAAAGCAATTTGTTGTT 60
QY 648 AGCGTGTGCTGACTGTTGAGGGGACGAATCAAGCATTTAGTTTGAATTC 707
DB 61 AGCGTGTGCTGACTGTTGAGGGGACGAATCAAGCATTTAGTTTGAATTC 120
QY 708 GATCTTAACATCAAGCCTGCTCATGAGAGAAAGACAGCAAGGCTTAAGTCAAAATCA 767
DB 121 GATCTTAACATCAAGCCTGCTCATGAGAGAAAGACAGCAAGGCTTAAGTCAAAATCA 180
QY 768 AAACATTGCTACTGATAGTGGCGGATGTCATTAACCTTGAGAAAGCTGACTACTA 827
DB 181 AAACATTGCTACTGATAGTGGCGGATGTCATTAACCTTGAGAAAGCTGACTACTA 240
QY 828 AAGGCTATTCAGAAACAAATTGATGGCTTAAGTCCCACTTAACGAGACTACTTTAGAGTC 887
DB 241 AAGGCTATTCAGAAACAAATTGATGGCTTAAGTCCCACTTAACGAGACTACTTTAGAGTC 300
QY 888 ATTGATTTTTCAGAGCAGTGCATTAAGTATGATGCAAAAGGCAAGGCTGACTTTGCTGAC 947

```

Db 301 ATTGATTTTGCAGCGATGCAACCATTAAGTCAAGAAACGCGAGGCTTACTTTGCTGAC 360
QY 948 AAAGATGTTGGTAACTTGGCGAGCCCACTGTCCAGAAATTTTGTAAAGGACAT 1007
Db 361 AAAGATGTTGGTAACTTGGCGAGCCCACTGTCCAGAAATTTTGTAAAGGACAT 420
QY 1008 GTGCGGTTAGACCATATAAAGAAAAACAATCAAAACCAAGGAAATCTGTGATGTG 1067
Db 421 GTGCGGTTAGACCATATAAAGAAAAACAATCAAAACCAAGGAAATCTGTGATGTG 480
QY 1068 GAATATATCTGTACATTTACTCCCTTAAACCTGTGATGACATTTTCAAGCCAGGCTCAAA 1127
Db 481 GAATATATCTGTACATTTACTCCCTTAAACCTGTGATGACATTTTCAAGCCAGGCTCAAA 540
QY 1128 GATCTAAGCTATTTGAAAAACATAGCATATGCTGACCAACATCATCTCAAGAAATTTCTA 1187
Db 541 GATCTAAGCTATTTGAAAAACATAGCATATGCTGACCAACATCATCTCAAGAAATTTCTA 600
QY 1188 GCTCAAGCAAAAGCAATTTTAAACAAAACCAACCAAGGCTATAGATTTTAAAGCTGAC 1247
Db 601 GCTCAAGCAAAAGCAATTTTAAACAAAACCAACCAAGGCTATAGATTTTAAAGCTGAC 660
QY 1248 TCCTCAATGCTCACTGATGACATGACATTTTCCGTAAGATTTTAAAGGATCAAGAG 1307
Db 661 TCCTCAATGCTCACTGATGACATGACATTTTCCGTAAGATTTTAAAGGATCAAGAG 720
QY 1308 TTTACTTACCGGTGTAAATTCGGGAAACAGCTTATAGATCATTAATAAATCTGTGTG 1367
Db 721 TTTACTTACCGGTGTAAATTCGGGAAACAGCTTATAGATCATTAATAAATCTGTGTG 780
QY 1368 AATGAAGAATAAACAACATGACCTGATCTGAGAAATTTACGTCCTTAAAAAAGG 1427
Db 781 AATGAAGAATAAACAACATGACCTGATCTGAGAAATTTACGTCCTTAAAAAAGG 840
QY 1428 GAAAGCCGTATGATCCCTTGTATGCGAGTCACTTGAACCTGTTACCATCAAAATGCTT 1487
Db 841 GAAAGCCGTATGATCCCTTGTATGCGAGTCACTTGAACCTGTTACCATCAAAATGCTT 900
QY 1488 GATGTGATATCAACCAATTTGCTTAAAGAGGAGCTCTTAAAGCTAGGAAAGCTAC 1547
Db 901 GATGTGATATCAACCAATTTGCTTAAAGAGGAGCTCTTAAAGCTAGGAAAGCTAC 960
QY 1548 TTAGACTTGAAGATTTATAGATCCCTGTAAGGCTAACTACTCAAAATGCTC 1607
Db 961 TTAGACTTGAAGATTTATAGATCCCTGTAAGGCTAACTACTCAAAATGCTC 1020
QY 1608 GATGCTTTGGTATTTATGACTATACCTTAACCTGGAAGAGTAAATCAGATGAC 1667
Db 1021 GATGCTTTGGTATTTATGACTATACCTTAACCTGGAAGAGTAAATCAGATGAC 1080
QY 1668 ACCAACCCTATCAATACCGTTTATATGAGGCAAGGAGCCGGAAGAGAGATGCTAGTAC 1727
Db 1081 ACCAACCCTATCAATACCGTTTATATGAGGCAAGGAGCCGGAAGAGAGATGCTAGTAC 1140
QY 1728 CATTTAGCTGTGCT 1742
Db 1141 CATTTAGCTGTGCT 1155

```

RESULT 4
US-08-568-393B-2
Sequence 2, Application US/08568393B
Patent No. 5876999

GENERAL INFORMATION:
APPLICANT: Hua-Lin Wu
APPLICANT: Huey-Yueh Shi
TITLE OF INVENTION: Preparation of novel streptokinase
TITLE OF INVENTION: mutants as improved thrombolytic agents
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jeing & Chang
STREET: Two No. 5876999th Second Street, Suite 290
CITY: San Jose

```

STATE: California
COUNTRY: USA
ZIP: 95113
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44MB
MEDIUM TYPE: Storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 6.1 on Window 3.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/568,393B
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Chi-Ping Chang
REGISTRATION NUMBER: 37,798
REFERENCE/DOCKET NUMBER:
TELEPHONE: (408) 288-8585
TELEFAX: (408) 288-8386
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1242 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: N
HYPOTHETICAL: N
ANTI-SENSE: N
FEATURE:
NAME/KEY: SK-K59E
LOCATION: DNA sequence No. 5876999174 and 175 have been changed
LOCATION: from AA to GG, and PROTEIN sequence No. 587699959 has been changed
LOCATION: from Lys to Glu.
OTHER INFORMATION:
US-08-568-393B-2
Query Match 54.4%; Score 1140.6; DB 2; Length 1242;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1146; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```

```

QY 588 ATTGCTGACCTGAGTGTGCTGACCGTCAATCTGTCAACCAAGCAATTTGTTGTT 647
Db 1 ATTGCTGACCTGAGTGTGCTGACCGTCAATCTGTCAACCAAGCAATTTGTTGTT 60
QY 648 AGCGTGTGCTGATCTGTTGAGGAGCAATCAAGCATTAATTTTGAATTC 707
Db 61 AGCGTGTGCTGATCTGTTGAGGAGCAATCAAGCATTAATTTTGAATTC 120
QY 708 GATCTAATCAACGACCTGCTCATGAGAGAAAGAGAGCAAGGCTTAAGTCCAAATCA 767
Db 121 GATCTAATCAACGACCTGCTCATGAGAGAAAGAGAGCAAGGCTTAAGTCCAAATCA 180
QY 121 GATCTAATCAACGACCTGCTCATGAGAGAAAGAGAGCAAGGCTTAAGTCCAAATCA 180
Db 768 AAACATTTGCTACTGATAGTGGCGAGTGCATTAACCTGGAAGAGTACTTACTA 827
QY 827 AAACATTTGCTACTGATAGTGGCGAGTGCATTAACCTGGAAGAGTACTTACTA 827
Db 181 AAACATTTGCTACTGATAGTGGCGAGTGCATTAACCTGGAAGAGTACTTACTA 240
QY 828 AAGCTATTCAGAAACATTTGATGCTTAACGTCACAGTAAAGCACTTGTAGAGTC 887
Db 241 AAGCTATTCAGAAACATTTGATGCTTAACGTCACAGTAAAGCACTTGTAGAGTC 300
QY 888 ATTGATTTTGCAGAGGAGTCAACCTTCTGATGCAAAACGGAAGTCTACTTGTGAC 947
Db 301 ATTGATTTTGCAGAGGAGTCAACCTTCTGATGCAAAACGGAAGTCTACTTGTGAC 360
QY 948 AAAGATGTTGGTAACTTGGCGAGCCCACTGTCCAGAAATTTTGTAAAGGACAT 1007
Db 361 AAAGATGTTGGTAACTTGGCGAGCCCACTGTCCAGAAATTTTGTAAAGGACAT 420
QY 1008 GTGCGGTTAGACCATATAAAGAAAAACAATCAAAACCAAGGAAATCTGTGATGTG 1067
Db 421 GTGCGGTTAGACCATATAAAGAAAAACAATCAAAACCAAGGAAATCTGTGATGTG 480

```

QY 1068 GAATATATCTGACAGTTTACTCCCTTAACCTGATGACGATTTCAAGCAGGCTCAAA 1127
DB 481 GAAATATCTGACAGTTTACTCCCTTAACCTGATGACGATTTCAAGCAGGCTCAAA 540
QY 1128 GATATCTGACAGTTTACTCCCTTAACCTGATGACGATTTCAAGCAGGCTCAAA 1187
DB 541 GATATCTGACAGTTTACTCCCTTAACCTGATGACGATTTCAAGCAGGCTCAAA 600
QY 1188 GCTGACAGCAAGAGATTTTAAACAAAACCAAGGCTATACGATTTATGACGATGAC 1247
DB 601 GCTGACAGCAAGAGATTTTAAACAAAACCAAGGCTATACGATTTATGACGATGAC 660
QY 1248 TCCCTCATCTGACATGACATGACATGATTTCCGATGATTTTCAATGATGACGATGAC 1307
DB 661 TCCCTCATCTGACATGACATGACATGATTTCCGATGATTTTCAATGATGACGATGAC 720
QY 1308 TTTACTTACCGGTGTTTAAATCCGGAACAGCTTATAGATTCATTAATTAATTAATTAATTAAT 1367
DB 721 TTTACTTACCGGTGTTTAAATCCGGAACAGCTTATAGATTCATTAATTAATTAATTAATTAAT 780
QY 1368 AATGACAGCAAGAGATTTTAAACAAAACCAAGGCTATACGATTTATGACGATGAC 1427
DB 781 AATGACAGCAAGAGATTTTAAACAAAACCAAGGCTATACGATTTATGACGATGAC 840
QY 1428 GAAAGCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1487
DB 841 GAAAGCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
QY 1488 GAT 1547
DB 901 GAT 960
QY 1548 TTAGACTTACAGAGATTTTATAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1607
DB 961 TTAGACTTACAGAGATTTTATAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
QY 1608 GAT 1667
DB 1021 GAT 1080
QY 1668 ACCAAGCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1727
DB 1081 ACCAAGCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
QY 1728 CATTTAGCTGATGAT 1742
DB 1141 CATTTAGCTGATGAT 1155

RESULT 5
US-08-488-940-19
Sequence 19, Application US/08488940
Patent No. 5854049
GENERAL INFORMATION:
APPLICANT: Reed, Guy L.
TITLE OF INVENTION: PLASMIN-RESISTANT STREPTOKINASE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,940
FILING DATE: 09-JUN-1995
CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 34,819
REFERENCE/DOCKET NUMBER: 05433/009001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 2566 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-488-940-19

Query Match 54.0%; Score 1132.4; DB 2; Length 2566;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1148; Conservative 0; Mismatches 6; Indels 2; Gaps 1;

QY 587 TATGCTGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 646
DB 896 TATGCTGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 955
QY 647 TAGCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 706
DB 956 TAGCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1015
QY 707 CGATCTAACATGACGACCTGCTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 766
DB 1016 CGATCTAACATGACGACCTGCTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 1073
QY 767 AAAACATTTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 826
DB 1074 AAAACATTTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1133
QY 827 AAAAGCTATTAAGAT 886
DB 1134 AAAAGCTATTAAGAT 1193
QY 887 CATGATTTTGAAGAT 946
DB 1194 CATGATTTTGAAGAT 1253
QY 947 CAAAGATGTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1006
DB 1254 CAAAGATGTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1313
QY 1007 TGTGCGGCTTACAGATTAAGAT 1066
DB 1314 TGTGCGGCTTACAGATTAAGAT 1373
QY 1067 GGAATATCTGACAGTTTACTCCCTTAACCTGATGACGATTTTCAAGCAGGCTCAAA 1126
DB 1374 GGAATATCTGACAGTTTACTCCCTTAACCTGATGACGATTTTCAAGCAGGCTCAAA 1433
QY 1127 AGATATCTGATTAAGAT 1186
DB 1434 AGATATCTGATTAAGAT 1493
QY 1187 AGCTAAGACAAAGAGATTTTAAACAAAACCAAGGCTATACGATTTATGACGATGAC 1246
DB 1494 AGCTAAGACAAAGAGATTTTAAACAAAACCAAGGCTATACGATTTATGACGATGAC 1553
QY 1247 CTCCTCAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1306
DB 1554 CTCCTCAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1613
QY 1307 GTTTACTTACCGGTGTTTAAATCCGGAACAGCTTATAGATTCATTAATTAATTAATTAATTAAT 1366
DB 1614 GTTTACTTACCGGTGTTTAAATCCGGAACAGCTTATAGATTCATTAATTAATTAATTAATTAAT 1673

OY	1367	GAATGAAGAAATTAACACACACGACCTGATCTGTGAGAAATATTAACGCTTTAAAAAAGG	1426
Db	1674	GAATGAAGAAATTAACACACACGACCTGATCTGTGAGAAATATTAACGCTTTAAAAAAGG	1733
OY	1427	GGAAAAAGCCGTATGATGCCCTTTGATGCGAGTCACCTTGAACTGTTCAACCATCAATAAGT	1486
Db	1724	GGAAAAAGCCGTATGATGCCCTTTGATGCGAGTCACCTTGAACTGTTCAACCATCAATAAGT	1793
OY	1487	TGATGTGCATATACCAACGAATTGCTTAAAAATGACGACGCTCTTTAACAGCTAGCGAACGTAA	1546
Db	1794	TGATGTGCATATACCAACGAATTGCTTAAAAATGACGACGCTCTTTAACAGCTAGCGAACGTAA	1853
OY	1547	CTTAGACTTCAGAGATTTATATACGATCCCTGCGATGATAGGCTTAACTCTAACCAATCT	1606
Db	1854	CTTAGACTTCAGAGATTTATATACGATCCCTGCGATGATAGGCTTAACTCTAACCAATCT	1913
OY	1607	CGATGCTTTTGGTATATATGACATATACCTTAACTGGAAGAATGACGATTAATCAAGATGA	1666
Db	1914	CGATGCTTTTGGTATATATGACATATACCTTAACTGGAAGAATGACGATTAATCAAGATGA	1973
OY	1667	CACCAACCGTATCATTAACCGTTTATATGAGGCAAGGACCCGAAGGAGAAATGCTAGCTA	1726
Db	1974	CACCAACCGTATCATTAACCGTTTATATGAGGCAAGGACCCGAAGGAGAAATGCTAGCTA	2033
OY	1727	CCATTTAGCTGTGTGT	1742
Db	2034	TCATTTAGCTTATGAT	2049

RESULT 6
 US-07-854-596B-42
 Sequence 42 Application US/07854596B
 Patent No. 5434073
 GENERAL INFORMATION:
 APPLICANT: Dawson, Keith M
 APPLICANT: Dawson, Michael G
 APPLICANT: Czaplewski, Lloyd G
 TITLE OF INVENTION: Proteins and nucleic acids
 NUMBER OF SEQUENCES: 73
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Dr. John J. McDonnell
 STREET: Ten South Wacker Drive, Suite. 3000
 CITY: Chicago
 STATE: IL
 COUNTRY: USA
 ZIP: 60606
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/854,596B
 FILING DATE: 03-JUN-1992
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: McDonnell, John J
 REGISTRATION NUMBER: 26,949
 REFERENCE/DOCKET NUMBER: 92,337
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312-715-1000
 TELEFAX: 312-715-1234
 TELEX: 910-221-5117
 INFORMATION FOR SEQ ID NO: 42:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1458 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 FEATURES:
 NAME/KEY: misc feature
 LOCATION: 1..1458

```

? OTHER INFORMATION: /note= "HindIII-streptokinase
? OTHER INFORMATION: fusion linked by Factor Xa cleavable IEGR"
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 1..1449
? FEATURE:
? NAME/KEY: mat_peptide
? LOCATION: 1..1449
? US-07-854-596B-42

```

Query Match	53.7%;	Score 1125.4;	DB 2;	Length 1458;
Best Local Similarity	97.8%;	Pred. No. 0;		
Matches 1141; Conservative	0;	Mismatches 26;	Indels 0;	Gaps 0;

QY	576	ACGAGTGTGGTATTTGCTGGAGACTGGAGGCGTGGCTAGACACGGTCACTGTGCAACAACACG	635
Db	196	ATCGAAGGTGAATTTGCTGGAGCTGGAGTGGCTGTAAGCCGTTCATCTGTCAACAACACG	255
QY	636	CAATTGGTGTGTACCGTTGCTGGTACTGTAAGGAGGAGCAATTCAGA CATTAGTCTTAAA	695
Db	256	CAATTAGTGTGTACCGTTGCTGGTACTGTAAGGAGGAGCAATTCAGA CATTAGTCTTAAA	315
QY	696	TTTTTTGAAATCGATCTTAACATCAAGACTGCTCATGAGGAAAGACAGACAGGCTTAA	755
Db	316	TTTTTTGAAATCGATCTTAACATCAAGACTGCTCATGAGGAAAGACAGACAGGCTTAA	375
QY	756	AGTCCAAATTCAAAACCATTTGGTACTGATAGTGGCGGAGTGTCAATTAACCTTGAGAAA	815
Db	376	AGTCCAAATTCAAAACCATTTGGTACTGATAGTGGCGGAGTGTCAATTAACCTTGAGAAA	435
QY	816	GCTGACTTACATAAAGGCTATTCAAGAACATTTGATCGCTAACGTCACAGTACAGACAC	875
Db	436	GCTGACTTACATAAAGGCTATTCAAGAACATTTGATCGCTAACGTCACAGTACAGACAC	495
QY	876	TACTTTGAGGTCAATTGATTTTGGCAAGCATGCAACATTACTGATCGAAACGGCAAGTTC	935
Db	496	TACTTTGAGGTCAATTGATTTTGGCAAGCATGCAACATTACTGATCGAAACGGCAAGTTC	555
QY	936	TACTTTGCTGCAAAAGATGTTCCGGTAACTTTGCCAGCCCAACCTGTGCOAAGAAATTTTGG	995
Db	556	TACTTTGCTGCAAAAGATGTTCCGGTAACTTTGCCAGCCCAACCTGTGCOAAGAAATTTTGG	615
QY	996	CTAAGCCGGAACATGTCGCGCTTAAAGCCATATTAAGAAAACCAATTCAAAACCAACGAAA	1055
Db	616	CTAAGCCGGAACATGTCGCGCTTAAAGCCATATTAAGAAAACCAATTCAAAACGAAA	675
QY	1056	TCTGTGATGTGGAATTAATACTGTACAGTTTACTCCCTTAAACCCTGATGACGATTTGAGA	1115
Db	676	TCTGTGATGTGGAATTAATACTGTACAGTTTACTCCCTTAAACCCTGATGACGATTTGAGA	735
QY	1116	CCAGGTCTCAAAAGATCTAAGCTATGTAACCACTAAGTATCGGTGACACATCAATCT	1175
Db	736	CCAGGTCTCAAAAGATCTAAGCTATGTAACCACTAAGTATCGGTGACACATCAATCT	795
QY	1176	CAAGAACTTACTAGCTCAAGGCACAAAAGCATTTTAAACAAAACCAACCCATCCAGGCTATACGATT	1235
Db	796	CAAGAACTTACTAGCTCAAGGCACAAAAGCATTTTAAACAAAACCCATCCAGGCTATACGATT	855
QY	1236	TATGAACTGTACTCCTCTCAATGTCACTCATAGCAATTTTCCGTACAGATTTTAAACA	1295
Db	856	TATGAACTGTACTCCTCTCAATGTCACTCATAGCAATTTTCCGTACAGATTTTAAACA	915
QY	1296	ATGATCAAGAAGTTTACTTACCGTGTAAAAATCGGAAACAAGCTTATAGATCAATATAA	1355
Db	916	ATGATCAAGAAGTTTACTTACCGTGTAAAAATCGGAAACAAGCTTATAGATCAATATAA	975
QY	1356	AAATCTGTGCTGAATGAAGAAATATAACAACCTGACCTGATCTCGAAGAAATATTACGTC	1415
Db	976	AAATCTGTGCTGAATGAAGAAATATAACAACCTGACCTGATCTCGAAGAAATATTACGTC	1035
QY	1416	CTTAAAAAAGGGGAAAAACCGGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTACAC	1475
Db	1036	CTTAAAAAAGGGGAAAAACCGGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTACAC	1095

QY 1476 ATCAATAGTGTGATGTCGATACCAAGAAATGCTTAAAGAGAGAGCTTTAAACAGCT 1535
 Db 1096 ATCAATAGTGTGATGTCGATACCAAGAAATGCTTAAAGAGAGAGCTTTAAACAGCT 1155
 QY 1536 AGCGAAGCTTACTTGAAGCTTCAAGATTTATATACGATCTCTGATTAAGGCTTAACTATC 1595
 Db 1156 AGCGAAGCTTACTTGAAGCTTCAAGATTTATATACGATCTCTGATTAAGGCTTAACTATC 1215
 QY 1596 TACAACATCTGAGCTTTGGTATATGAGCTATACCTTAAGTGAAGAGAT 1655
 Db 1216 TACAACATCTGAGCTTTGGTATATGAGCTATACCTTAAGTGAAGAGAT 1275
 QY 1656 AATCAGATGACCAACCGTATCAATACCGTTATATGAGGCAAGGCAAGAGAG 1715
 Db 1276 AATCAGATGACCAACCGTATCAATACCGTTATATGAGGCAAGGCAAGAGAG 1335
 QY 1716 AATGCTAGTACCAATTTAGCTGAGTGT 1742
 Db 1336 AATGCTAGTACCAATTTAGCTGAGTGT 1362
 RESULT 7
 US-07-703-778D-1
 ; Sequence 1, Application US/07703778D
 ; Patent No. 5296366
 ; GENERAL INFORMATION:
 ; APPLICANT: Garcia, M.P.E. et al
 ; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND EXPRESSION
 ; TITLE OF INVENTION: OF A GENE WHICH CODES FOR STREPTOKINASE
 ; NUMBER OF SEQUENCES: 14
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Stenger, Michaelson, Spivak and Wallace, Esq.
 ; STREET: Parkway 109 Office Center, 328 Newman Springs Road,
 ; STREET: P. O. Box 8489
 ; CITY: Red Bank
 ; STATE: New Jersey
 ; COUNTRY: USA
 ; ZIP: 07701
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3 1/2" 1.44mb IBM compatible diskette
 ; COMPUTER: IBM PS/2 Model 80
 ; OPERATING SYSTEM: MS-DOS 5.0
 ; SOFTWARE: Microsoft Word for Windows 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07703,778D
 ; FILING DATE: 19910522
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Michaelson, Peter L.
 ; REGISTRATION NUMBER: 30090
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (908) 530-6671
 ; TELEFAX: (908) 530-6584
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1245 base pairs
 ; TYPE: NUCLEOTIDE WITH CORRESPONDING PROTEIN
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: genomic DNA
 ; ORIGINAL SOURCE:
 ; ORGANISM: Streptococcus equisimilis from group C of Lanfield
 ; ORGANISM: definition
 ; IMMEDIATE SOURCE: ATCC-9542 strain
 ; FEATURE: from 1 to 1245 bp mature peptide
 ; OTHER INFORMATION:
 ; OTHER INFORMATION: Properties: Streptokinase gene
 ; OTHER INFORMATION: The gene product binds to human plasminogen
 ; OTHER INFORMATION: The gene product is an activator of human plasminogen
 ; US-07-703-778D-1

Query Match 53.7%; Score 1124.6; DB 2; Length 1245;
 Best Local Similarity 98.4%; Pred. No. 0;
 Matches 1136; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
 QY 588 ATTGCTGAGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACCAAGCCAAATGGTGT 647
 Db 1 ATTGCTGAGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACCAAGCCAAATGGTGT 60
 QY 648 AGCGTCTGCTGATCTGTGAGGGGACGAATCAAGCATTTAGCTTAAATTTTGAATTC 707
 Db 61 AGCGTCTGCTGATCTGTGAGGGGACGAATCAAGCATTTAGCTTAAATTTTGAATTC 120
 QY 708 GATCTAACATCAGACCGTGCATGAGGAAGACAGCAAGGCTTAAGTCAAAATCA 767
 Db 121 GATCTAACATCAGACCGTGCATGAGGAAGACAGCAAGGCTTAAGTCAAAATCA 180
 QY 768 AAACCATTTGCTACTGATAGTGGCGCATGTCAATAACTTGAAGAAAGTGACTTACTA 827
 Db 181 AAACCATTTGCTACTGATAGTGGCGCATGTCAATAACTTGAAGAAAGTGACTTACTA 240
 QY 828 AAGGCTATTCAAGAACATTTGATGCTTAACGTCAAGTAAAGAGTAACTTTGAGGTC 887
 Db 241 AAGGCTATTCAAGAACATTTGATGCTTAACGTCAAGTAAAGAGTAACTTTGAGGTC 300
 QY 888 ATTGATTTTCAAGCGCATGCAACCATTAATGATGGAAGACGCAAGTCTACTTGTGAC 947
 Db 301 ATTGATTTTCAAGCGCATGCAACCATTAATGATGGAAGACGCAAGTCTACTTGTGAC 360
 QY 948 AAAGATGTTGCGTAACTTGGCGCAACCACTGTCCAAAGATTTTGTGAAGGAGAT 1007
 Db 361 AAAGATGTTGCGTAACTTGGCGCAACCACTGTCCAAAGATTTTGTGAAGGAGAT 420
 QY 1008 GTGGCGTTAGACCAATTAAGAAAGAAACCAATCAAAACCAAGGAAATCTGTGATGTG 1067
 Db 421 GTGGCGTTAGACCAATTAAGAAAGAAACCAATCAAAACCAAGGAAATCTGTGATGTG 480
 QY 1068 GAATATATCTGATCAAGTTACTTCCCTTAAACCTGATGACATTTCAAGACGAGTCTCAA 1127
 Db 481 GAATATATCTGATCAAGTTACTTCCCTTAAACCTGATGACATTTCAAGACGAGTCTCAA 540
 QY 1128 GATACCTAAGCTAATTAAGAAACCTAGTATGAGTGAACCAATCAATCAAGAAATTTCTA 1187
 Db 541 GATACCTAAGCTAATTAAGAAACCTAGTATGAGTGAACCAATCAATCAAGAAATTTCTA 600
 QY 1188 GCTCAAGCAAAAGCAATTTTAAACAAACCAACCAAGCTATACGATTTTGAAGCTGAC 1247
 Db 601 GCTCAAGCAAAAGCAATTTTAAACAAACCAACCAAGCTATACGATTTTGAAGCTGAC 660
 QY 1248 TCCTCAATCGTCACTGACATGACATGACATTTTCCGTACGATTTTCAATGATCAAGAG 1307
 Db 661 TCCTCAATCGTCACTGACATGACATGACATTTTCCGTACGATTTTCAATGATCAAGAG 720
 QY 1308 TTACTTACCGTGTAAATAATCGGGAACAAGCTATAGATCAATTAATAATCGGTCTG 1367
 Db 721 TTACTTACCGTGTAAATAATCGGGAACAAGCTATAGATCAATTAATAATCGGTCTG 780
 QY 1368 AATGAAGAATTAACCAACCTGACCTGATCTGAGAAATTTTCTTAAATAAAGG 1427
 Db 781 AATGAAGAATTAACCAACCTGACCTGATCTGAGAAATTTTCTTAAATAAAGG 840
 QY 1428 GAAAGCCGTATGATCCCTTTGATGCGAGTCACTTGAAGCTGTTCAACATCAATACGTT 1487
 Db 841 GAAAGCCGTATGATCCCTTTGATGCGAGTCACTTGAAGCTGTTCAACATCAATACGTT 900
 QY 1488 GATGTGATTCACAGCAATTTGCTTAAAGAGAGAGCTTTTACAGCTAGAGGATAC 1547
 Db 901 GATGTGATTCACAGCAATTTGCTTAAAGAGAGAGCTTTTACAGCTAGAGGATAC 960
 QY 1548 TTAGACTTGAGAGATTTATGATCCCGATTAAGGCTTAACTTCACTCAACATCTC 1607
 Db 961 TTAGACTTGAGAGATTTATGATCTCTGTATAGGCTTAACTTCACTCAACATCTC 1020

Qy	1608	GATGCTTTGGAAATTAAGCACTAACCTTAAC	CTGGAAGAAAGAGAAATACAGATAC	1667
Db	1021	GATGCTTTTGGATTAATGACTATACCTTAAC	CTGGAAGAAAGAGAAATACAGATAC	1080
Qy	1668	ACCAACCGTATCATPACCGTTTATATGAGCA	CCCGAAGAGAAATGCTAGCTAC	1727
Db	1081	ACCAACCGTATCATPACCGTTTATATGAGCA	CCCGAAGAGAAATGCTAGCTAT	1140
Qy	1728	CATTAGCTGGTGGT	1742	
Db	1141	CATTAGCTGATGAT	1155	

RESULT 8

```

US-07-854-596B-14
Sequence 14, Application US/07854596B
Patent No. 5434073

GENERAL INFORMATION:
APPLICANT: Dawson, Keith M
APPLICANT: Hunter, Michael G
APPLICANT: Caplawski, Lloyd G
TITLE OF INVENTION: Proteins and nucleic acids
NUMBER OF SEQUENCES: 73
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dr. John J. McDonnell
STREET: Ten South Wacker Drive, Suite 3000
CITY: Chicago
STATE: IL
COUNTRY: USA
ZIP: 60606

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/854,596B
FILING DATE: 03-JUN-1992
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: McDonnell, John J
REGISTRATION NUMBER: 26,949
REFERENCE/DOCKET NUMBER: 92,337
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
TELEX: 910-221-5317

INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 1335 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: misc feature
LOCATION: 1..1335
OTHER INFORMATION: /note="Streptokinase gene from S
OTHER INFORMATION: equisimilis"
FEATURE:
NAME/KEY: CDS
LOCATION: 7..1326
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 7..1326
US-07-854-596B-14

```

Query Match	53.5%	Score 1122.4;	DB 2;	Length 1335;
Best Local Similarity	98.2%;	Pred. No. 0;		
Matches 1135;	Conservative 0;	Mismatches 21;	Indels 0;	Gaps 0;

587 TATTGCTGCACTGAGTGGCTGCTAGACCGTTCATCTGTCAACAACGCCAATTGGTGT 646

Db	84	TATTCGTGAGCACTGATGCGCTGCTAAGACCGTTCATCTGTCAACAACAGCCAAATTAGTTGT	143
QY	647	TAGCGTTCGTGTGTACTGTGTAAGGCGGACGAATCAAGAACAATTAGTCTTAAATTTTTGAAAT	706
Db	144	TAGCGTTCGTGTCTGTGTGTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTGAAAT	203
QY	707	CGATCTTAACTACAGACCTTGCTCATGTGAGGGAAGAACAAGACGAAGGCTTAAACTCCAAAATC	766
Db	204	TGACCTTAACTACAGACCTTGCTCATGTGAGGGAAGAACAAGACGAAGGCTTAAACTCCAAAATC	263
QY	767	AAAACACTTTTCTCTGTGATAGTGGCGCATGTCAATTAATTGAGAAAAGCTGACTTA	826
Db	264	AAAACACTTTTCTCTGTGATAGTGGCGCATGTCAATTAATTGAGAAAAGCTGACTTA	323
QY	827	AAAGGCTATTCAAGAACCAATTGATCGCTAAACGTCACAAGTAACGACGTACTTTGAGGT	886
Db	324	AAAGGCTATTCAAGAACCAATTGATCGCTAAACGTCACAAGTAACGACGTACTTTGAGGT	383
QY	887	CATTGATTTTTCGAAGCCGATGCAACCTTATCTGATTCGAAACGGCAAGGCTTACTTTGCTGA	946
Db	384	CATTGATTTTTCGAAGCCGATGCAACCTTATCTGATTCGAAACGGCAAGGCTTACTTTGCTGA	443
QY	947	CAAAAGATGCTCGGGAACCTTGCGGACCCCAACCTGTCCAAAGATTTTTTGCTAAGCGGACA	1006
Db	444	CAAAAGATGCTCGGGAACCTTGCGGACCCCAACCTGTCCAAAGATTTTTTGCTAAGCGGACA	503
QY	1007	TGTGCGCGTTTGAACCATATAAAGAAAAACAATACAAAACCAAGCGAAATCTGTTGATGT	1066
Db	504	TGTGCGCGTTTGAACCATATAAAGAAAAACAATACAAAATCAAGCGAAATCTGTTGATGT	563
QY	1067	GGAATATATCTGTACAGTTTACTCCCTTAAACCTGTATGACGATTTCAACCAAGCTTCA	1126
Db	564	GGAATATATCTGTACAGTTTACTCCCTTAAACCTGTATGACGATTTCAACCAAGCTTCA	623
QY	1127	AGATACCTAAGCTATTGAAAAACATAGAGTATGGGTGACACCATGCAATCTGCAAGAAATTACT	1186
Db	624	AGATACCTAAGCTATTGAAAAACATAGAGTATGGGTGACACCATGCAATCTGCAAGAAATTACT	683
QY	1187	AGCTCAAGCACAAAAGCAATTTTAAACAAAACCAACCAGGCTATACGATTTATGAAACGTGA	1246
Db	684	AGCTCAAGCACAAAAGCAATTTTAAACAAAACCAACCAGGCTATACGATTTATGAAACGTGA	743
QY	1247	CTCTCAATCGTCACTCATGCAATGACATTTTCCGTACGATTTTACCAATGATCAAGA	1306
Db	744	CTCTCAATCGTCACTCATGCAATGACATTTTCCGTACGATTTTACCAATGATCAAGA	803
QY	1307	GTTTACTTAAACCGTATTAAAAATCGGGAAACAAGCTTATAGATTCAAATAAAAAATCTGCTCT	1366
Db	804	GTTTACTTAAACCGTATTAAAAATCGGGAAACAAGCTTATAGATTCAAATAAAAAATCTGCTCT	863
QY	1367	GAATGAAGAAATTAACAACAACCTGACCTGATCTCTGAAGAAATATTACGCTCTTAAAAAAG	1426
Db	864	GAATGAAGAAATTAACAACAACCTGACCTGATCTCTGAAGAAATATTACGCTCTTAAAAAAG	923
QY	1427	GGAAGAGCCGTATGATCCCTTGTGATCGCAGTCACTTGAAACTGTTCACCAATCAATAACT	1486
Db	924	GGAAGAGCCGTATGATCCCTTGTGATCGCAGTCACTTGAAACTGTTCACCAATCAATAACT	983
QY	1487	TGATGTGATATCCAACGAATTTGCTTAAAAAGTGACACACTCTTAAACAGTATAGCGAACCTAA	1546
Db	984	TGATGTGATATCCAACGAATTTGCTTAAAAAGTGACACACTCTTAAACAGTATAGCGAACCTAA	1043
QY	1547	CTTAGACCTTGAAGATTTATACGATTCCTCGGTATAGGCTTAAACTACTCTTCAACAATCT	1606
Db	1044	CTTAGACCTTGAAGATTTATACGATTCCTCGGTATAGGCTTAAACTACTCTTCAACAATCT	1103
QY	1607	CGATGCTTTTGTATATATAGACTATACCTTAACTGAGAAAAGTAGAGATTAATCAAGATGA	1666
Db	1104	CGATGCTTTTGTATATATAGACTATACCTTAACTGAGAAAAGTAGAGATTAATCAAGATGA	1163
QY	1667	CACCAACCGTATCATTAACCGTTTATATATGGGAACGACCCGGAAGAGAGAAATGCTTAGCTA	1726
Db	1164	CACCAACCGTATCATTAACCGTTTATATATGGGAACGACCCGGAAGAGAGAAATGCTTAGCTA	1223

US-07-854-596B-34
; Sequence 34, Application US/07854596B
; Patent No. 5434073
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith M
; APPLICANT: Hunter, Michael G
; APPLICANT: Caplewski, Lloyd G
; TITLE OF INVENTION: Proteins and nucleic acids
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. John J. McDonnell
; STREET: Ten South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/854,596B
; FILING DATE: 03-JUN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,337
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELETYPE: 910-221-5317
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2589 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..2589
; OTHER INFORMATION: /note=
; OTHER INFORMATION: "OmpA-Streptokinase-streptokinase fusion linked
; OTHER INFORMATION: by thrombin-cleavable VEGFvrbg"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 4..2580
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 4..2580
; US-07-854-596B-34
Query Match 53.5%; Score 1122.2; DB 2; Length 2589;
Best Local Similarity 98.0%; Pred. No. 0;
Matches 1136; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
QY 584 TCGTATGTCGACCTGAGTGGCTGTAACCGTCATCTGTCAACAACAGCCAAATGTGT 643
DB 1335 TCGAATGTGTGACCTGAGTGGCTGTAACCGTCATCTGTCAACAACAGCCAAATGTGT 1394
QY 644 TGTTCACGTCGTCGTGTAAGTGGAGGGAAGAAATCAAGACATTAAGTCTTAATTTTGA 703
DB 1395 TGTTCACGTCGTCGTGTAAGTGGAGGGAAGAAATCAAGACATTAAGTCTTAATTTTGA 1454
QY 704 AATGATCTTAACATCAAGACCTGCTCATGAGAGAAAGACAGCAAGCCTTAAGTCCAA 763
DB 1455 AATTAACCTTAACATCAAGACCTGCTCATGAGAGAAAGACAGCAAGCCTTAAGTCCAA 1514
QY 764 ATCAAAACCATTTGCTACTGATAGTGGCGAGTGTCAATTAACCTTGAAGAAAGCTGACT 823
DB 1515 ATCAAAACCATTTGCTACTGATAGTGGCGAGTGTCAATTAACCTTGAAGAAAGCTGACT 1574

QY 824 ACTTAAAGCTATTCAAGAACATTTGATGCTTAACGTCACAGTAACGACACTACTTTGA 883
DB 1575 ACTTAAAGCTATTCAAGAACATTTGATGCTTAACGTCACAGTAACGACACTACTTTGA 1634
QY 884 GGTTCATTGATTTTGAACGAGATGCAACCATTAAGTGAAGAAAGGCTTAAGTTC 943
DB 1635 GGTTCATTGATTTTGAACGAGATGCAACCATTAAGTGAAGAAAGGCTTAAGTTC 1694
QY 944 TGACAAAGATGCTTCGTTAACCTTGGACCCCAACCTGTCTCAAGATTTTGTCAAGCG 1003
DB 1695 TGACAAAGATGCTTCGTTAACCTTGGACCCCAACCTGTCTCAAGATTTTGTCAAGCG 1754
QY 1004 ACATGTCGCGTTAGACCATATTAAGAAACCAATCAAAACCAAGGAAATCTGTTGA 1063
DB 1755 ACATGTCGCGTTAGACCATATTAAGAAACCAATCAAAACCAAGGAAATCTGTTGA 1814
QY 1064 TGTGATATTAAGTATGTAAGTTTACTCCCTTAAACCTGATGACGATTTGACAGCGTCT 1123
DB 1815 TGTGATATTAAGTATGTAAGTTTACTCCCTTAAACCTGATGACGATTTGACAGCGTCT 1874
QY 1124 CAAAGATTAAGCTATTGAAACACTAGCTATGCGTGACACATCAATCTCAAGATT 1183
DB 1875 CAAAGATTAAGCTATTGAAACACTAGCTATGCGTGACACATCAATCTCAAGATT 1934
QY 1184 ACTAGCTCAGCAAAAGCAATTTTAAACAAACCAACCGGCTATGCAATTTATGAACG 1243
DB 1935 ACTAGCTCAGCAAAAGCAATTTTAAACAAACCAACCGGCTATGCAATTTATGAACG 1994
QY 1244 TGAATCTCAATGCTCATGATGACATGATGATGATGATGATGATGATGATGATGATG 1303
DB 1995 TGAATCTCAATGCTCATGATGACATGATGATGATGATGATGATGATGATGATGATG 2054
QY 1304 AGAGTTTACTTACCGTGTAAATAATGCGAAACAAAGCTTATGATGATGATGATGATG 1363
DB 2055 AGAGTTTACTTACCGTGTAAATAATGCGAAACAAAGCTTATGATGATGATGATGATG 2114
QY 1364 TCTGAATGAAGAAATTAACAACCTGACGATCTGACGAAATTAATGCTTAAATA 1423
DB 2115 TCTGAATGAAGAAATTAACAACCTGACGATCTGACGAAATTAATGCTTAAATA 2174
QY 1424 AGGGGAAAGCCGATATGATCCCTTGTGACAGTCACTTGAACCTGTTCACATCAATA 1483
DB 2175 AGGGGAAAGCCGATATGATCCCTTGTGACAGTCACTTGAACCTGTTCACATCAATA 2234
QY 1484 CGTTGATGTCATCAACCAAGATTTGCTAAAGTGAAGCAAGCTTAAACAGTACGAACG 1543
DB 2235 CGTTGATGTCATCAACCAAGATTTGCTAAAGTGAAGCAAGCTTAAACAGTACGAACG 2294
QY 1544 TAACTTGAATCTTCAAGATTTATGATCTCTGATTAAGGCTTAACTACTTCAACAA 1603
DB 2295 TAACTTGAATCTTCAAGATTTATGATCTCTGATTAAGGCTTAACTACTTCAACAA 2354
QY 1604 TCTGATGCTTTTGTGATTAAGACTATGATCTTAACTTAACTGAAAGTAGAGTAATCA 1663
DB 2355 TCTGATGCTTTTGTGATTAAGACTATGATCTTAACTTAACTGAAAGTAGAGTAATCA 2414
QY 1664 TGAACCAACCGATATCAACCGTTTATATGAGGCAAGCAACCGAAGAGAGAAATGCTAG 1723
DB 2415 TGAACCAACCGATATCAACCGTTTATATGAGGCAAGCAACCGAAGAGAGAAATGCTAG 2474
QY 1724 CTATCATTATGCTGCTGCT 1742
DB 2475 CTATCATTATGCTGCTGCT 2493
RESULT 11
US-07-854-596B-25
; Sequence 25, Application US/07854596B
; Patent No. 5434073
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith M
; APPLICANT: Hunter, Michael G

APPLICANT: Czaplowski, Lloyd G
TITLE OF INVENTION: Proteins and nucleic acids
NUMBER OF SEQUENCES: 73
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dr. John J. McDonnell
STREET: Ten South Wacker Drive, Suite 3000
CITY: Chicago
STATE: IL
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/854,596B
FILING DATE: 03-JUN-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: McDonnell, John J
REGISTRATION NUMBER: 26,949
REFERENCE/DOCKET NUMBER: 92,337
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
TELEX: 910-221-5317
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 1257 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..1257
OTHER INFORMATION: /note="Methionyl-streptokinase"
FEATURE:
NAME/KEY: CDS
LOCATION: 4..1248
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 4..1248
US-07-854-596B-25
Query Match 53.5%; Score 1121.4; DB 2; Length 1257;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 1134; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 588 ATTGCTGACCTGAGTGGCTGCTGACCGTCCATCTGTCAACACAGCCCAATGTTGTT 647
DB 7 ATTGCTGACCGTGAAGTGGCTGCTGACCGTCCATCTGTCAACACAGCCCAATGTTGTT 66
QY 648 AGGTTGCTGCTGCTGCTGAGGGGACGATCAAGACATTAGTCTTAATTTTGAATC 707
DB 67 AGGTTGCTGCTGCTGCTGAGGGGACGATCAAGACATTAGTCTTAATTTTGAATC 126
QY 708 GATCTAACATCAGACCTGCTCATGAGAGAAAGACAGAGCAAGGCTTAAGTCAAAATCA 767
DB 127 GACCTAACATCAGACCTGCTCATGAGAGAAAGACAGAGCAAGGCTTAAGTCAAAATCA 186
QY 768 AAACCATTTGCTACTGATAGTGGCGCGATGTCACATTAACCTTGAGAAAGTGAATTACTA 827
DB 187 AAACCATTTGCTACTGATAGTGGCGCGATGTCACATTAACCTTGAGAAAGTGAATTACTA 246
QY 828 AAGGCTATTCAAGAAATTTGATTCGCTTAAGTCAACGTAACGACGACTTCTTGAAGTC 887
DB 247 AAGGCTATTCAAGAAATTTGATTCGCTTAAGTCAACGTAACGACGACTTCTTGAAGTC 306
QY 888 ATTGATTTTGAAGGATGCAACCATTAATGATGGAAGGCAAGGCTACTTGTGCTGAC 947

DB 307 ATTGATTTTGAAGGATGCAACCATTAATGATGGAAGGCAAGGCTACTTGTGCTGAC 366
QY 948 AAAGATGTTGGTGAACCTTGCCGACCCAACTGTCAGAAATTTTGTAGCCGACAT 1007
DB 367 AAAGATGTTGGTGAACCTTGCCGACCCAACTGTCAGAAATTTTGTAGCCGACAT 426
QY 1008 GTGCCGTTAGACATATTAAGAAAGAAACCAATACAAACCAAGCAATCTGTGATG 1067
DB 427 GTGCCGTTAGACATATTAAGAAAGAAACCAATACAAACCAAGCAATCTGTGATG 486
QY 1068 GAATATCTGTACAGTTTACTCTCTTAACCTGTATGACGATTTGACAGAGTCTCAA 1127
DB 487 GAATATCTGTACAGTTTACTCTCTTAACCTGTATGACGATTTGACAGAGTCTCAA 546
QY 1128 GATATTAAGCTATTGAAAACTACTATCGGTGACACATGACATCTCAAGATTTACTA 1187
DB 547 GATATTAAGCTATTGAAAACTACTATCGGTGACACATGACATCTCAAGATTTACTA 606
QY 1188 GCTCAAGCACAAAGCATTTTAAACAAACCAAGGCTATGATTTATGACGTCAC 1247
DB 607 GCTCAAGCACAAAGCATTTTAAACAAACCAAGGCTATGATTTATGACGTCAC 666
QY 1248 TCTCAATGTCATCTCATGACATGACATTTTCCGTAGCATTTTACAAATGATCAAG 1307
DB 667 TCTCAATGTCATCTCATGACATGACATTTTCCGTAGCATTTTACAAATGATCAAG 726
QY 1308 TTTACTTACCGGTGTTAAAAATCGGGAACAAGTTATAGATTAATTAATAATCTGTCTG 1367
DB 727 TTTACTTACCGGTGTTAAAAATCGGGAACAAGTTATAGATTAATTAATAATCTGTCTG 786
QY 1368 AATGAAGAAATTAACCAACACTGACCTGATCTGAGAAATTTAGTCTTTAAAAAGGG 1427
DB 787 AATGAAGAAATTAACCAACACTGACCTGATCTGAGAAATTTAGTCTTTAAAAAGGG 846
QY 1428 GAAAAACCGTATGATCCCTTTGATCGACATGCTTAAGAACTGTTCAACATACGTT 1487
DB 847 GAAAAACCGTATGATCCCTTTGATCGACATGCTTAAGAACTGTTCAACATACGTT 906
QY 1488 GATGTCATCAACCAAGAAATTCCTTAAGAAAGTGAAGCTCTTAAACGTAAGCAATGTAAC 1547
DB 907 GATGTCATCAACCAAGAAATTCCTTAAGAAAGTGAAGCTCTTAAACGTAAGCAATGTAAC 966
QY 1548 TTAGACTCAGAGATTTATAGATCCCTGATTAAGGTTAAAGTCTTACAAACAATTC 1607
DB 967 TTAGACTCAGAGATTTATAGATCCCTGATTAAGGTTAAAGTCTTACAAACAATTC 1026
QY 1608 GATGCTTTTGTATTAAGACTATACCTTAAGTGAAGAAAGTGAAGATTAACGATGAC 1667
DB 1027 GATGCTTTTGTATTAAGACTATACCTTAAGTGAAGAAAGTGAAGATTAACGATGAC 1086
QY 1668 ACCAACCGTATCAATACCGTTTATATGAGGCAAGCGACCGAAGAGAGAAATGCTAGTAC 1727
DB 1087 ACCAACCGTATCAATACCGTTTATATGAGGCAAGCGACCGAAGAGAGAAATGCTAGTAC 1146
QY 1728 CATTAGCTGTGCT 1742
DB 1147 CATTAGCTGTGAT 1161

RESULT 12
US-07-854-596B-18
Sequence 18, Application US/07854596B
Patent No. 5434073
GENERAL INFORMATION:
APPLICANT: Dawson, Keith M
APPLICANT: Hunter, Michael G
APPLICANT: Czaplowski, Lloyd G
TITLE OF INVENTION: Proteins and nucleic acids
NUMBER OF SEQUENCES: 73
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dr. John J. McDonnell
STREET: Ten South Wacker Drive, Suite 3000
CITY: Chicago

STATE: IL
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/854,596B
FILING DATE: 03-JUN-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: McDowell, John J
REGISTRATION NUMBER: 26,949
REFERENCE/DOCKET NUMBER: 92,337
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
TELEX: 910-221-5317
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 1317 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: misc feature
LOCATION: 1..1317
OTHER INFORMATION: /note="OmpA fused to mature
FEATURE:
NAME/KEY: CDS
LOCATION: 4..1308
FEATURE:
NAME/KEY: mat peptide
LOCATION: 4..1308
US-07-854-596B-18

Query Match 53.5%; Score 1121.4; DB 2; Length 1317;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 1134; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
QY 588 ATTGCTGACCTGATGAGTGGCTGCTAGACCGTCCATCTGTCAACAACGCCAATGGTTGTT 647
DB 67 ATTGCTGACCTGATGAGTGGCTGCTAGACCGTCCATCTGTCAACAACGCCAATGGTTGTT 126
QY 648 AGCGTTGCTGTAAGTGTGAGGGGAGCAATCAAGACATTAGTCTTAATTTTGAATC 707
DB 127 AGCGTTGCTGTAAGTGTGAGGGGAGCAATCAAGACATTAGTCTTAATTTTGAATC 186
QY 708 GATCTACATCAAGCACTGCTCATGAGAGAAAGACAGAGCAAGCTTAAAGTCCAAATCA 767
DB 187 GATCTACATCAAGCACTGCTCATGAGAGAAAGACAGAGCAAGCTTAAAGTCCAAATCA 246
QY 768 AAACCAATTTGCTACTGATGATGAGCGCGGATGTCACATPAACTTGAGAAAGTGACTTA 827
DB 247 AAACCAATTTGCTACTGATGATGAGCGCGGATGTCACATPAACTTGAGAAAGTGACTTA 306
QY 828 AAGGCTATTCAGAACCAATGATGCTAAGTCCAGTCAAGTCAAGCACTTAAGTGAAGTC 887
DB 307 AAGGCTATTCAGAACCAATGATGCTAAGTCCAGTCAAGTCAAGCACTTAAGTGAAGTC 366
QY 888 ATTGATTTTGGCAAGGATGCAACCATTAAGTCAAGCAAGGCTTACTTGGTGAAC 947
DB 367 ATTGATTTTGGCAAGGATGCAACCATTAAGTCAAGCAAGGCTTACTTGGTGAAC 426
QY 948 AAAGATGTTGGTGAACCTTGGCGACCAACCTGTCGAAGATTTTGGTGAACGCAAT 1007
DB 427 AAAGATGTTGGTGAACCTTGGCGACCAACCTGTCGAAGATTTTGGTGAACGCAAT 486
QY 1008 GTGGCGGTTAGACCATATAAAGAAAAACCAATACAAACCAAGCAATCTGTGATGTC 1067

|||||
DB 487 GTGGCGGTTAGACCATATAAAGAAAAACCAATACAAATCAAGCAAGCAATCTGTGATGTC 546
QY 1068 GAATATATCTGATGACCTTATCCCTTAAACCTGATACGATTTTCAAGCCAGGCTGAAA 1127
DB 547 GAATATATCTGATGACCTTATCCCTTAAACCTGATACGATTTTCAAGCCAGGCTGAAA 606
QY 1128 GATCTAAGCTATTTGAAACCACTAGTATGAGTGAACCAATCAATCAATCAATCAATCA 1187
DB 607 GATCTAAGCTATTTGAAACCACTAGTATGAGTGAACCAATCAATCAATCAATCAATCA 666
QY 1188 GCTCAAGCAAGGATTTTAAACCAAAACCAACCAAGCTATTAAGATTTTAAAGCTGAC 1247
DB 667 GCTCAAGCAAGGATTTTAAACCAAAACCAACCAAGCTATTAAGATTTTAAAGCTGAC 726
QY 1248 TCCCTCAATCTGCTACTGATGACATGATGATTTTCCGTAAGATTTTAAACCAAGTCA 1307
DB 727 TCCCTCAATCTGCTACTGATGACATGATGATTTTCCGTAAGATTTTAAACCAAGTCA 786
QY 1308 TTTACTTACCGTGTAAATAATCGGAGCAAGCTTATGATCAATTAATAAATCTGCTG 1367
DB 787 TTTACTTACCGTGTAAATAATCGGAGCAAGCTTATGATCAATTAATAAATCTGCTG 846
QY 1368 AATGAAGAAATTAACCAACCACTGACCTGATCTGAGAAATATTAAGTCTTAAATAAGG 1427
DB 847 AATGAAGAAATTAACCAACCACTGACCTGATCTGAGAAATATTAAGTCTTAAATAAGG 906
QY 1428 GAAAAGCGTATGATCCCTTTGATCGCATCTTGAACCTGTAACCTGTTCAACATTAAGTT 1487
DB 907 GAAAAGCGTATGATCCCTTTGATCGCATCTTGAACCTGTAACCTGTTCAACATTAAGTT 966
QY 1488 GATGTGATACCAAGCAATGCTTAATAAGAGAGAGCTTAAAGCTAGAGCAAGCTGAC 1547
DB 967 GATGTGATACCAAGCAATGCTTAATAAGAGAGAGCTTAAAGCTAGAGCAAGCTGAC 1026
QY 1548 TTAGACTTCAGAGATTTATAGATCCCTGATTAAGGCTTAACTACTCAACATCTC 1607
DB 1027 TTAGACTTCAGAGATTTATAGATCCCTGATTAAGGCTTAACTACTCAACATCTC 1086
QY 1608 GATGCTTTTGTATATGACATATACCTTAACCTGAAAGTAGAGATTAACGATGAC 1667
DB 1087 GATGCTTTTGTATATGACATATACCTTAACCTGAAAGTAGAGATTAACGATGAC 1146
QY 1668 ACCAAGCTATTCATTAACCGTTTATATGAGGACGACCCGAGAGAGATGCTAGCTAC 1727
DB 1147 ACCAAGCTATTCATTAACCGTTTATATGAGGACGACCCGAGAGAGATGCTAGCTAT 1206
QY 1728 CATTTAGCTGTGAT 1742
DB 1207 CATTTAGCTGTGAT 1221
RESULT 13
US-07-854-596B-46
Sequence 46, Application US/07854596B
Patent No. 5434073
GENERAL INFORMATION:
APPLICANT: Dawson, Keith M
APPLICANT: Hunter, Michael G
APPLICANT: Czaplewski, Lloyd G
TITLE OF INVENTION: Proteins and nucleic acids
NUMBER OF SEQUENCES: 73
CORRESPONDENCE ADDRESSES:
STREET: Ten South Wacker Drive, Suite 3000
CITY: Chicago
STATE: IL
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

```

SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/854,596B
FILING DATE: 03-JUN-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: McDonnell, John J
REGISTRATION NUMBER: 26,949
REFERENCE/DOCKET NUMBER: 92,337
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
TELEX: 910-221-5317
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 1467 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: misc feature
LOCATION: 1..1467
OTHER INFORMATION: /note= "Streptokinase-hirudin
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1449
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 1..1449
US-07-854-596B-46

```

```

Query Match      53.5%; Score 1121.4; DB 2; Length 1467;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 1134; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 588 ATTGCTGACGCTGAGTGGCTGCTAGACCGTGCATCTGTCAACAACGCGAATTGGTTGT 647
DB 1 ATTGCTGACGCTGAGTGGCTGCTAGACCGTGCATCTGTCAACAACGCGAATTGGTTGT 60

QY 648 AGCGTTGCTGTACTGTGAGGGGAGCAATCAAGACTTAAATTTTGAATC 707
DB 61 AGCGTTGCTGTACTGTGAGGGGAGCAATCAAGACTTAAATTTTGAATC 120

QY 708 GATTTAACAATCAGACCTGCTCATGAGAAAGACAGAGCAAGCTTAAGTCCAAATCA 767
DB 121 GATTTAACAATCAGACCTGCTCATGAGAAAGACAGAGCAAGCTTAAGTCCAAATCA 180

QY 768 AAACCAATTTGCTAATGATGCGGAGATGCTCATTAATTTGAAAGCTGACTTA 827
DB 181 AAACCAATTTGCTAATGATGCGGAGATGCTCATTAATTTGAAAGCTGACTTA 240

QY 828 AAGGCTAATCAAGAAATTTGATGCTAAGTCCAGCACTAAGTCTTGAAGTC 887
DB 241 AAGGCTAATCAAGAAATTTGATGCTAAGTCCAGCACTAAGTCTTGAAGTC 300

QY 888 ATTGATTTTGAAGGATGCAACATTAAGTGAACGCAAGGCTTAAGTCTTGAAGTC 947
DB 301 ATTGATTTTGAAGGATGCAACATTAAGTGAACGCAAGGCTTAAGTCTTGAAGTC 360

QY 948 AAAAGATGTTGGGTAACCTTCCGAGACCACTGCTCAAGAAATTTTGAAGGAT 1007
DB 361 AAAAGATGTTGGGTAACCTTCCGAGACCACTGCTCAAGAAATTTTGAAGGAT 420

QY 1008 GTGCGGCTTGAACCATATAAGAAAAACAATAACAACGAAATGTTGAATG 1067
DB 421 GTGCGGCTTGAACCATATAAGAAAAACAATAACAACGAAATGTTGAATG 480

QY 1068 GAATATACGTACAGTTTACCTCCCTTAAACCTGATGAGATTTGAGACCAAGGCTCAAA 1127
DB 481 GAATATACGTACAGTTTACCTCCCTTAAACCTGATGAGATTTGAGACCAAGGCTCAAA 540

```

```

QY 1128 GATATCAAGCTATTGAAAAACATAGCTATCGGTGACACCATCATCTCAAGAAATTA 1187
DB 541 GATATCAAGCTATTGAAAAACATAGCTATCGGTGACACCATCATCTCAAGAAATTA 600

QY 1188 GCTAAGACCAAGCAATTTTAAACAAAAACAACCAAGGCTATAGCAATTAAGACGTGAC 1247
DB 601 GCTAAGACCAAGCAATTTTAAACAAAAACCATCAGGCTATAGCAATTAAGACGTGAC 660

QY 1248 TCCCAATTCGCTACATGACAAATGACATTTTCGTAGATTTTAAACCAATGATCAAGAG 1307
DB 661 TCCCAATTCGCTACATGACAAATGACATTTTCGTAGATTTTAAACCAATGATCAAGAG 720

QY 1308 TTTACTTACCGTGTAAAAAATCGGAAAACAAGCTTAAGATCAATTAATAAAATCGTCTG 1367
DB 721 TTTACTTACCGTGTAAAAAATCGGAAAACAAGCTTAAGATCAATTAATAAAATCGTCTG 780

QY 1368 AATGAGAAATTAACAACACTGACCTGATCTCTGAGAAATTAAGTCTTAAAAAAGG 1427
DB 781 AATGAGAAATTAACAACACTGACCTGATCTCTGAGAAATTAAGTCTTAAAAAAGG 840

QY 1428 GAAAAGCGTATGATTCCTTTGATCGAGTCACTTGAACCTTCAACATCAATACGT 1487
DB 841 GAAAAGCGTATGATTCCTTTGATCGAGTCACTTGAACCTTCAACATCAATACGT 900

QY 1488 GATGTCGATACCAAGCAATGCTTAAAAAGTAGAGAGCTTTAAAGAGTGAAGCAAGTAA 1547
DB 901 GATGTCGATACCAAGCAATGCTTAAAAAGTAGAGAGCTTTAAAGAGTGAAGCAAGTAA 960

QY 1548 TTAGACTTCAGAGATTTATAGATCTCGTATTAAGGCTTAACCTTCAACAAATCTC 1607
DB 961 TTAGACTTCAGAGATTTATAGATCTCGTATTAAGGCTTAACCTTCAACAAATCTC 1020

QY 1608 GATGCTTTGGTATTAAGACTATACCTTAACTGAAAAAGTAGAGATTAACAGATGAC 1667
DB 1021 GATGCTTTGGTATTAAGACTATACCTTAACTGAAAAAGTAGAGATTAACAGATGAC 1080

QY 1668 ACCAAGCGTATCAATTAACGTTTATATAGGGCAAGGACCCGAAGAGAAATGCTAGCTAC 1727
DB 1081 ACCAAGCGTATCAATTAACGTTTATATAGGGCAAGGACCCGAAGAGAAATGCTAGCTAT 1140

QY 1728 CATTAGCTGTGCT 1742
DB 1141 CATTAGCTGTGCT 1155

```

```

RESULT 14
US-09-374-038-13
Sequence 13, Application US/09374038
Patent No. 6309873
GENERAL INFORMATION:
APPLICANT: Madrazo, Isis Del Carmen Torrens
APPLICANT: Garcia, Jose De Jesus De la Fuente
APPLICANT: Ojalvo, Ariana Garcia
APPLICANT: Menendez, Ailana Seralena
APPLICANT: Escalona, Elder Pupo
APPLICANT: Masso, Julio Raul Fernandez
APPLICANT: Griego, Martha De Jesus Gonzalez
TITLE OF INVENTION: STREPTOKINASE MUTANTS
FILE REFERENCE: Sequence listings 1-14 re: 976-5
Patent No. 6309873
CURRENT APPLICATION NUMBER: US/09/374,038
CURRENT FILING DATE: 1999-08-13
NUMBER OF SEQ ID NOS: 14
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 13
LENGTH: 1122
TYPE: DNA
ORGANISM: Streptococcus equisimilis
US-09-374-038-13

```

```

Query Match      52.2%; Score 1093.6; DB 3; Length 1122;
Best Local Similarity 98.7%; Pred. No. 3,66-312;
Matches 1102; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

```

QY 588 ATTGCTGGAAGCTGAGTGGCTGAGACCGTCCATCTGTCAACAACAGCCAAATGGTGT 647
DB 4 ATTGCTGGAAGCTGAGTGGCTGAGACCGTCCATCTGTCAACAACAGCCAAATGGTGT 63
QY 648 AGCGTTGCTGTAAGTGGTGAAGGAGCAATCAAGACATTAAGTCTTAATTTTGAATC 707
DB 64 AGCGTTGCTGTAAGTGGTGAAGGAGCAATCAAGACATTAAGTCTTAATTTTGAATC 123
QY 708 GATCTTAACATCAAGACCTGCTCATGAGAGAAACAGAGCCTTAAGTCCAAATCA 767
DB 124 GACCTTAACATCAAGACCTGCTCATGAGAGAAACAGAGCCTTAAGTCCAAATCA 183
QY 768 AAACATTTGCTAGTGAATGAGCGGAGTGTCAATTAACCTTGAAGAGTGAATCA 827
DB 184 AAACATTTGCTAGTGAATGAGCGGAGTGTCAATTAACCTTGAAGAGTGAATCA 243
QY 828 AAGGCTATTCAAGAACATTTGATCGCTAACGTCACAGTAAACGACATTAATTGAGGTC 887
DB 244 AAGGCTATTCAAGAACATTTGATCGCTAACGTCACAGTAAACGACATTAATTGAGGTC 303
QY 888 ATTGAATTTGCAAGGATGCAACCATTAATGATGCAAGAGGCAAGGCTTAATTGCTGAC 947
DB 304 ATTGAATTTGCAAGGATGCAACCATTAATGATGCAAGAGGCAAGGCTTAATTGCTGAC 363
QY 948 AAAGATGCTGGTGAACCTTGGCCAGCCCAACCTGTCAGAAATTTTGTAGCCGACAT 1007
DB 364 AAAGATGCTGGTGAACCTTGGCCAGCCCAACCTGTCAGAAATTTTGTAGCCGACAT 423
QY 1008 GTGCGCGTTAGACCATATTAAGAAAAACCAATACAAACCAAGCAAAATCTGTTGATGTG 1067
DB 424 GTGCGCGTTAGACCATATTAAGAAAAACCAATACAAACCAAGCAAAATCTGTTGATGTG 483
QY 1068 GAATTAATCTGTAAGTTTAATCTCCCTTAAACCTGATGACGATTTGACACGAGTCTCAAA 1127
DB 484 GAATTAATCTGTAAGTTTAATCTCCCTTAAACCTGATGACGATTTGACACGAGTCTCAAA 543
QY 1128 GATTAAGCTATTAAGAAAAACATAGCTATCGGTGACACATCAATCTCAAGAAATTAATCA 1187
DB 544 GATTAAGCTATTAAGAAAAACATAGCTATCGGTGACACATCAATCTCAAGAAATTAATCA 603
QY 1188 GCTCAAGCAACAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTATGAAACGTGAC 1247
DB 604 GCTCAAGCAACAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTATGAAACGTGAC 663
QY 1248 TCCTCAATCTGCACTCATGACCAATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 1307
DB 664 TCCTCAATCTGCACTCATGACCAATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 723
QY 1308 TTTTACTTACCGTGTAAATAATGGGAAACAGCTTATAGATCAATTAATAATCTGCTG 1367
DB 724 TTTTACTTACCGTGTAAATAATGGGAAACAGCTTATAGATCAATTAATAATCTGCTG 783
QY 1368 AATGAAGAAATTAACAACATGACCTGATCTCTGAGAAATATTAAGCTCTTAAATAAGAG 1427
DB 784 AATGAAGAAATTAACAACATGACCTGATCTCTGAGAAATATTAAGCTCTTAAATAAGAG 843
QY 1428 GAAAAGCCGTATGATCCCTTTGATGCACTGCACTTGAACCTGTTCAATCAATAGCTT 1487
DB 844 GAAAAGCCGTATGATCCCTTTGATGCACTGCACTTGAACCTGTTCAATCAATAGCTT 903
QY 1488 GATGTCGATCAACAAGATTTGCTTAAATAATGAGCAGCTTAAACGCTTACCGCAAGCTAAC 1547
DB 904 GATGTCGATCAACAAGATTTGCTTAAATAATGAGCAGCTTAAACGCTTACCGCAAGCTAAC 963
QY 1548 TTAAGCTTCAAGATTTTATAGATCTGATGAAGGCTTAACTACTTCAACATCTC 1607
DB 964 TTAAGCTTCAAGATTTTATAGATCTGATGAAGGCTTAACTACTTCAACATCTC 1023
QY 1608 GATGCTTTTGTATTAAGATTAATCTTAATCTGAAAGAGAGATTAATCAAGATGAC 1667
DB 1024 GATGCTTTTGTATTAAGATTAATCTTAATCTGAAAGAGAGATTAATCAAGATGAC 1083

QY 1668 ACCAACCGTATCATTAACCGTTTATATGGCAAGCA 1703
DB 1084 ACCAACCGTATCATTAACCGTTTATATGGCAAGCA 1119
RESULT 15
US-09-658-179-13
; Sequence 13, Application US/09658179
; Patent No. 6413759
; GENERAL INFORMATION:
; APPLICANT: Madrazo, Isis Del Carmen Torrens
; APPLICANT: Garcia, Jose De Jesus De La Fuente
; APPLICANT: Ojalvo, Ariana Garcia
; APPLICANT: Menendez, Alina Seralena
; APPLICANT: Escalona, Eider Pupo
; APPLICANT: Masso, Julio Raul Fernandez
; APPLICANT: Griego, Martha De Jesus Gonzalez
; TITLE OF INVENTION: STREPTOKINASE MUTANTS
; FILE REFERENCE: Sequence Listings 1-14 re: 976-5
; Patent No. 6413759
; CURRENT APPLICATION NUMBER: US/09/658,179
; CURRENT FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 1122
; TYPE: DNA
; ORGANISM: Streptococcus equisimilis
US-09-658-179-13
Query Match 52.2%; Score 1093.6; DB 3; Length 1122;
Best Local Similarity 98.7%; Pred. No. 3.6e-312;
Matches 1102; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 588 ATTGCTGGAAGCTGAGTGGCTGAGACCGTCCATCTGTCAACAACAGCCAAATGGTGT 647
DB 4 ATTGCTGGAAGCTGAGTGGCTGAGACCGTCCATCTGTCAACAACAGCCAAATGGTGT 63
QY 648 AGCGTTGCTGTAAGTGGTGAAGGAGCAATCAAGACATTAAGTCTTAATTTTGAATC 707
DB 64 AGCGTTGCTGTAAGTGGTGAAGGAGCAATCAAGACATTAAGTCTTAATTTTGAATC 123
QY 708 GATCTTAACATCAAGACCTGCTCATGAGAGAAACAGAGCCTTAAGTCCAAATCA 767
DB 124 GACCTTAACATCAAGACCTGCTCATGAGAGAAACAGAGCCTTAAGTCCAAATCA 183
QY 768 AAACATTTGCTAGTGAATGAGCGGAGTGTCAATTAACCTTGAAGAGTGAATCA 827
DB 184 AAACATTTGCTAGTGAATGAGCGGAGTGTCAATTAACCTTGAAGAGTGAATCA 243
QY 828 AAGGCTATTCAAGAACATTTGATCGCTAACGTCACAGTAAACGACATTAATTGAGGTC 887
DB 244 AAGGCTATTCAAGAACATTTGATCGCTAACGTCACAGTAAACGACATTAATTGAGGTC 303
QY 888 ATTGAATTTGCAAGGATGCAACCATTAATGATGCAAGAGGCAAGGCTTAATTGCTGAC 947
DB 304 ATTGAATTTGCAAGGATGCAACCATTAATGATGCAAGAGGCAAGGCTTAATTGCTGAC 363
QY 948 AAGGATGCTGGTGAACCTTGGCCAGCCCAACCTGTCAGAAATTTTGTAGCCGACAT 1007
DB 364 AAGGATGCTGGTGAACCTTGGCCAGCCCAACCTGTCAGAAATTTTGTAGCCGACAT 423
QY 1008 GTGCGCGTTAGACCATATTAAGAAAAACCAATACAAACCAAGCAAAATCTGTTGATGTG 1067
DB 424 GTGCGCGTTAGACCATATTAAGAAAAACCAATACAAACCAAGCAAAATCTGTTGATGTG 483
QY 1068 GAATTAATCTGTAAGTTTAATCTCCCTTAAACCTGATGACGATTTTCAACAGGCTCAAA 1127
DB 484 GAATTAATCTGTAAGTTTAATCTCCCTTAAACCTGATGACGATTTTCAACAGGCTCAAA 543
QY 1128 GATTAAGCTATTAAGAAAAACATAGCTATCGGTGACACATCAATCTCAAGAAATTAATCA 1187
DB 544 GATTAAGCTATTAAGAAAAACATAGCTATCGGTGACACATCAATCTCAAGAAATTAATCA 603

```
QY 1188 GCTCAAGCACAAAGCATTTTAAACAAAACCAAGGCTATACGATTTATGAACGTGAC 1247
DB 604 GCTCAAGCACAAAGCATTTTAAACAAAACCAAGGCTATACGATTTATGAACGTGAC 663
QY 1248 TCCTCAATCGTCATCTGATGACATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 1307
DB 664 TCCTCAATCGTCATCTGATGACATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 723
QY 1308 TTTACTTACCGTGTAAATAATCGGGAACAAGCTTATAGATCAATPAAAAAATCGTCTG 1367
DB 724 TTTACTTACCGTGTAAATAATCGGGAACAAGCTTATAGATCAATPAAAAAATCGTCTG 783
QY 1368 AATGAAGAAAATAACAACTGACCTGATCTGTGAAATATACTGTCCTTAAAAAAGG 1427
DB 784 AATGAAGAAAATAACAACTGACCTGATCTGTGAAATATACTGTCCTTAAAAAAGG 843
QY 1428 GAAAAAGCGTATGATCCCTTTGATGCGATGCACTTGAAACTGTTCAACCATCAATACGTT 1487
DB 844 GAAAAAGCGTATGATCCCTTTGATGCGATGCACTTGAAACTGTTCAACCATCAATACGTT 903
QY 1488 GATGTGATACCAAGAAATGCTPAAAAAGTAGAGAGCTCTTAACAGCTAGGGAACGTAC 1547
DB 904 GATGTGATACCAAGAAATGCTPAAAAAGTAGAGAGCTCTTAACAGCTAGGGAACGTAC 963
QY 1548 TTAGACTTCAGAGATTTATACGATCTGTGATPAGGCTAACTACTTACACACATCTC 1607
DB 964 TTAGACTTCAGAGATTTATACGATCTGTGATPAGGCTAACTACTTACACACATCTC 1023
QY 1608 GATGCTTTGGTATTTATGACTATACCTTAACTGAAAAAGTAGAGATTAATCACGATGAC 1667
DB 1024 GATGCTTTGGTATTTATGACTATACCTTAACTGAAAAAGTAGAGATTAATCACGATGAC 1083
QY 1668 ACCAACCCTATCAATACCGTTTATATAGGCGAAGCGA 1703
DB 1084 ACCAACCCTATCAATACCGTTTATATAGGCGAAGCGA 1119
```

Search completed: February 1, 2006, 12:43:44
Job time : 371.056 secs


```

Db      121  GGGTCTCAACGACAGAGACGATCATGCGCACCCGTGGCCAGAGACCCAAACGCTGCGCG 180
Qy      181  AGATCGATCCCGCGCAAAATTAATACAGACTCATATGGAGAGACCAACGGTTCCCTC 240
Db      181  AGATCGATCCCGCGCAAAATTAATACAGACTCATATGGAGAGACCAACGGTTCCCTC 240
Qy      241  TGAATAATATTTTGTAACTTAAGAGAGATATACATGGTGAAGCAACAGATT 300
Db      241  TGAATAATATTTTGTAACTTAAGAGAGATATACATGGTGAAGCAACAGATT 300
Qy      301  GTAACCATAGCTGAAGAAGTGTGATCATGCTGCTGGACCTTCTATGTGTGGAGAA 360
Db      301  GTAACCATAGCTGAAGAAGTGTGATCATGCTGCTGGACCTTCTATGTGTGGAGAA 360
Qy      361  AGGTGGAGAAAGGACCGGACGATCATCTTGCACTTTAGAAATAGATGACATCAG 420
Db      361  AGGTGGAGAAAGGACCGGACGATCATCTTGCACTTTAGAAATAGATGACATCAG 420
Qy      421  ACAAGAGACATCTTAAGAAATGAGACACTGAGACAAAGAGATATGAGGAACC 480
Db      421  ACAAGAGACATCTTAAGAAATGAGACACTGAGACAAAGAGATATGAGGAACC 480
Qy      481  TGCTCCAGTGCATCTGCAAGGCAAGCGGAGAGAGTGAAGTGAAGGACACACT 540
Db      481  TGCTCCAGTGCATCTGCAAGGCAAGCGGAGAGAGTGAAGTGAAGGACACACT 540
Qy      541  CTGTGACAGACCATAGAGCGGATCTGGCCCCCTTACCGATGTTGATTTGTCGAC 600
Db      541  CTGTGACAGACCATAGAGCGGATCTGGCCCCCTTACCGATGTTGATTTGTCGAC 600
Qy      601  AGTGGCTGTAGACCGTGCATCTGTCAACAAAGCCAAATGTGTTAGGGTGTGCTGA 660
Db      601  AGTGGCTGTAGACCGTGCATCTGTCAACAAAGCCAAATGTGTTAGGGTGTGCTGA 660
Qy      661  CTGTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTGAATGATCTAACATCAC 720
Db      661  CTGTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTGAATGATCTAACATCAC 720
Qy      721  GACCTGCTCATGAGGAGAAAGACAGACCAAGGCTTAAGTCCAAATCAAAATCTTA 780
Db      721  GACCTGCTCATGAGGAGAAAGACAGACCAAGGCTTAAGTCCAAATCAAAATCTTA 780
Qy      781  CTGATAGTGGCGGATGTCATTAACCTTGAAGAGCTGAAGCTTAAGCTTAAGCTTA 840
Db      781  CTGATAGTGGCGGATGTCATTAACCTTGAAGAGCTGAAGCTTAAGCTTAAGCTTA 840
Qy      841  AACATTTGATCGCTAACGTCACAGTAAAGACGACTCTTGAAGTCAATGATTTTCAA 900
Db      841  AACATTTGATCGCTAACGTCACAGTAAAGACGACTCTTGAAGTCAATGATTTTCAA 900
Qy      901  GCGATGCAACCATTAAGTGAAGCGCAAGGCTTACTTGTGCTGACAAAGATGTTCCG 960
Db      901  GCGATGCAACCATTAAGTGAAGCGCAAGGCTTACTTGTGCTGACAAAGATGTTCCG 960
Qy      961  TAACTTGGCCGACCAACCTGTCCAGAAATTTTGTCAACGACGATGCGCGCTTGAC 1020
Db      961  TAACTTGGCCGACCAACCTGTCCAGAAATTTTGTCAACGACGATGCGCGCTTGAC 1020
Qy      1021  CATATTAAGAAAACCAATACAAAACCAAGCAAAATCTGTGATGTGAATATATCTGAC 1080
Db      1021  CATATTAAGAAAACCAATACAAAACCAAGCAAAATCTGTGATGTGAATATATCTGAC 1080
Qy      1081  AGTTTACTCCCTTAAACCTGATGAGATTTGACCAAGGCTCAAAAGATATAGCTAT 1140
Db      1081  AGTTTACTCCCTTAAACCTGATGAGATTTGACCAAGGCTCAAAAGATATAGCTAT 1140
Qy      1141  TGAAGAACCTAGTATGGTGAACATCATCTCAAGAAATTAAGTCTCAAGCAAAA 1200
Db      1141  TGAAGAACCTAGTATGGTGAACATCATCTCAAGAAATTAAGTCTCAAGCAAAA 1200
Qy      1201  GGATTTTAAACAAAACCAACCGGCTATGATTTATGAACGATCTCCGATGCTCA 1260

```

```

Db      1201  GCATTTTAAACAAAACCAACCGGCTATAGCATTTATGAACGTGACTCTCAATGCTCA 1260
Qy      1261  CTCAATGACAAATGACATTTTCCGTACGATTTTAAACCAATGATCAAGTTTACTCCGG 1320
Db      1261  CTCAATGACAAATGACATTTTCCGTACGATTTTAAACCAATGATCAAGTTTACTCCGG 1320
Qy      1321  TTAATAATCGGGAACAAGCTTATAGATCAATTAATAAATCTGCTCAATGAAGAAATTA 1380
Db      1321  TTAATAATCGGGAACAAGCTTATAGATCAATTAATAAATCTGCTCAATGAAGAAATTA 1380
Qy      1381  ACAACACTGACCTGATCTTGAAGAAATATAGTCTTAAATAAAGGGGAAAACCGTATG 1440
Db      1381  ACAACACTGACCTGATCTTGAAGAAATATAGTCTTAAATAAAGGGGAAAACCGTATG 1440
Qy      1441  ATCCCTTGAATCGAGTCACTTGAACCTGTTACCATCAAAATACCTTGAATGTCATCA 1500
Db      1441  ATCCCTTGAATCGAGTCACTTGAACCTGTTACCATCAAAATACCTTGAATGTCATCA 1500
Qy      1501  ACGAATTCCTAAATAAGTGAAGGCTCTTAAACAGTGAAGGAAAGTAACTTGAATCAAG 1560
Db      1501  ACGAATTCCTAAATAAGTGAAGGCTCTTAAACAGTGAAGGAAAGTAACTTGAATCAAG 1560
Qy      1561  ATTTATAGATCTCTGATTAAGGCTTAAACTTCTACAAATCTGATGCTTTTGGTA 1620
Db      1561  ATTTATAGATCTCTGATTAAGGCTTAAACTTCTACAAATCTGATGCTTTTGGTA 1620
Qy      1621  TTATGACATTAATCTTAACTGGAAGATGAAGATATCAAGATGACACCAACCGTATCA 1680
Db      1621  TTATGACATTAATCTTAACTGGAAGATGAAGATATCAAGATGACACCAACCGTATCA 1680
Qy      1681  TTAACGTTTATATGGGCAAGGACCCGAAAGAGAAAGTAACTAGTAACTTAAAGCTGG 1740
Db      1681  TTAACGTTTATATGGGCAAGGACCCGAAAGAGAAAGTAACTAGTAACTTAAAGCTGG 1740
Qy      1741  GTGCGCAGGCGCAACAGATTTGACCATAGCTGAGAAAGTGTGTTGATCATGCTGGGA 1800
Db      1741  GTGCGCAGGCGCAACAGATTTGACCATAGCTGAGAAAGTGTGTTGATCATGCTGGGA 1800
Qy      1801  CTTCTATGTGTGTGAGAAACGTGGGAAAGCCCTTACCAAGGCTGATGATGATTA 1860
Db      1801  CTTCTATGTGTGTGAGAAACGTGGGAAAGCCCTTACCAAGGCTGATGATGATTA 1860
Qy      1861  GTACTGCTGTGGGAAAGGCGGAGGACATCACTGTGACTTGAATATGATGATCAACG 1920
Db      1861  GTACTGCTGTGGGAAAGGCGGAGGACATCACTGTGACTTGAATATGATGATCAACG 1920
Qy      1921  ATCAGACACAAAGACATCTATGAATTTGAGACACTGAGCAAGAGATTAATCGAG 1980
Db      1921  ATCAGACACAAAGACATCTATGAATTTGAGACACTGAGCAAGAGATTAATCGAG 1980
Qy      1981  GAAACTGCTCAAGTGCATCTGCAACAGGCAACGCGGAGAGAGTGAAGTGAAGGC 2040
Db      1981  GAAACTGCTCAAGTGCATCTGCAACAGGCAACGCGGAGAGAGTGAAGTGAAGGC 2040
Qy      2041  ACACCTGTGACACCAACATGAGCGGATCTGGCCCTTCCAGATGTTGGTTAG 2096
Db      2041  ACACCTGTGACACCAACATGAGCGGATCTGGCCCTTCCAGATGTTGGTTAG 2096

RESULT 2
US-09-940-235-11
; Sequence 11, Application US/09940235
; Publication No. US20030059921A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammarra
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION

```


APPLICANT: Sahni, Girish
APPLICANT: Roy, Chait
APPLICANT: Rajagopal, Kamnara
APPLICANT: Nihalani, Deepak
APPLICANT: Sundaram, Vasudha
APPLICANT: Yadav, Mahavir
TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
TITLE OF INVENTION: PROTEIN
FILE REFERENCE: 07064-009002
CURRENT APPLICATION NUMBER: US/09/940,235
CURRENT FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: 09/471,349
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: IN 3825/DEL/98
PRIOR FILING DATE: 1998-12-24
NUMBER OF SEQ ID NOS: 28
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 9
LENGTH: 1541
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Hybrid cassette
US-09-940-235-9

Query Match 71.4%; Score 1496.2; DB 3; Length 1541;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1501; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 588 ATTGCTGACCTGAGGCGTGTGACGCTCCATCTGTCAACAACAGCCAAATGTTGTT 647
DB 33 AATGCTGCGCTGGAATGCGTACTAGATCGCTCTGTAATAACGCCAAATGTTGTT 92
QY 648 AGCGTTGCTGACTGTGTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTGAATC 707
DB 93 AGCGTTGCTGACTGTGTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTGAATC 152
QY 708 GATCTAACATCAAGACCTGCTCATGAGGAAAGACAGAGCAAGCTTAACTCCAAATCA 767
DB 153 GATCTAACATCAAGACCTGCTCATGAGGAAAGACAGAGCAAGCTTAACTCCAAATCA 212
QY 768 AAACCATTTGCTACTGATAGTGGCGGATGTCATPAACTTTGAGAAAGTGAATTACTA 827
DB 213 AAACCATTTGCTACTGATAGTGGCGGATGTCATPAACTTTGAGAAAGTGAATTACTA 272
QY 828 AAGGCTATTCAAGAACATTTGATCGCTAACGTCACAGTAAAGACGACTTCTTGAAGTC 887
DB 273 AAGGCTATTCAAGAACATTTGATCGCTAACGTCACAGTAAAGACGACTTCTTGAAGTC 332
QY 888 ATTGATTTTGGCAAGCGATCAACCATTTACTGATGAAACGGCAAGCTTACTTTGCTGAC 947
DB 333 ATTGATTTTGGCAAGCGATCAACCATTTACTGATGAAACGGCAAGCTTACTTTGCTGAC 392
QY 948 AAAGATGTTGCGTAACTTTGCCGACCCAACTGTCCAGAAATTTTGTCTAAGCGGACAT 1007
DB 333 AAAGATGTTGCGTAACTTTGCCGACCCAACTGTCCAGAAATTTTGTCTAAGCGGACAT 452
QY 1008 GTGGCGGTTAGACCATATTAAGAAAAACAATACAAACCAACGGAATCTGTTGATGTC 1067
DB 453 GTGGCGGTTAGACCATATTAAGAAAAACAATACAAACCAACGGAATCTGTTGATGTC 512
QY 1068 GAATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTGAGACAGGCTCTCAA 1127
DB 513 GAATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTGAGACAGGCTCTCAA 572
QY 1128 GATCTAAGCTATTTGAAAACTAGCTATCGGTGACACATCACTCTCAAGAAATTAATA 1187
DB 573 GATCTAAGCTATTTGAAAACTAGCTATCGGTGACACATCACTCTCAAGAAATTAATA 632
QY 1188 GGTCAAGCAAGAACTTTTAAACAAAAACCCGAGGCTATACGATTTATGAACGTCAC 1247

DB 633 GCTCAAGCAAGAACTTTTAAACAAAAACCCAGGCTATACGATTTATGAACGTGAC 692
QY 1248 TCCCTCAATCGCTCACTGACGACAAATGACATTTCCGTACGATTTTAAACAAATGACCAAG 1307
DB 693 TCCCTCAATCGCTCACTGACGACAAATGACATTTCCGTACGATTTTAAACAAATGACCAAG 752
QY 1308 TTTACTTACCGGTGTTAAAAATCCGGAAACAGCTTATAGATCAATPAAAAATCTGTCTG 1367
DB 753 TTTACTTACCGGTGTTAAAAATCCGGAAACAGCTTATAGATCAATPAAAAATCTGTCTG 812
QY 1368 AATGAAAGAAATPAAACCACTGACCTGATCTCTGAGAAATTAACGTCCTTAAAAAAGG 1427
DB 813 AATGAAAGAAATPAAACCACTGACCTGATCTCTGAGAAATTAACGTCCTTAAAAAAGG 872
QY 1428 GAAAAAGCGTATGATCCCTTTGATCGACATCTTGAACCTGTTACCATTAATAGTT 1487
DB 873 GAAAAAGCGTATGATCCCTTTGATCGACATCTTGAACCTGTTACCATTAATAGTT 932
QY 1488 GATGTGATACCAAGAAATGCTPAAAAAGTGAAGAGCTTTAAACAGCTAGGAAAGTAAAC 1547
DB 933 GATGTGATACCAAGAAATGCTPAAAAAGTGAAGAGCTTTAAACAGCTAGGAAAGTAAAC 992
QY 1548 TTGACTTCAAGATTTTATACGATCTCTGATTAAGCTTAACTTCAACAAATCTC 1607
DB 993 TTGACTTCAAGATTTTATACGATCTCTGATTAAGCTTAACTTCAACAAATCTC 1052
QY 1608 GATGCTTTTGTATTTATGACATTAACCTTAACTGGAAGAAATGAGATTAATCAGATGAC 1667
DB 1053 GATGCTTTTGTATTTATGACATTAACCTTAACTGGAAGAAATGAGATTAATCAGATGAC 1112
QY 1668 ACCAAGCGTATCAATACCGTTTATATGAGGAAAGGACCCGAAAGAGAAATGCTAGCTAC 1727
DB 1113 ACCAAGCGTATCAATACCGTTTATATGAGGAAAGGACCCGAAAGAGAAATGCTAGCTAC 1172
QY 1728 CATTTAGCTGTGTGCGCCAGGCGCAACAGATTTGATCCATAGCTGAGAAATGTTTAT 1787
DB 1173 CATTTAGCTGTGTGCGCCAGGCGCAACAGATTTGATCCATAGCTGAGAAATGTTTAT 1232
QY 1788 CATGCTGTGGAACCTTCTATGTTGTCGAGAAACGTGGAGAAACCTTACCAAGGCTGG 1847
DB 1233 CATGCTGTGGAACCTTCTATGTTGTCGAGAAACGTGGAGAAACCTTACCAAGGCTGG 1292
QY 1848 ATGATGTGATATTTGATCTTCCCTGGGAGAAAGGACCGCAACCTGACCTTCTAGA 1907
DB 1293 ATGATGTGATATTTGATCTTCCCTGGGAGAAAGGACCGCAACCTGACCTTCTAGA 1352
QY 1908 AATAGATGCAAGCATCAGACACAAGGACATCTTATAGAAATTTGAGACAACCTGAGAGAAG 1967
DB 1353 AATAGATGCAAGCATCAGACACAAGGACATCTTATAGAAATTTGAGAGACAACCTGAGAGAAG 1412
QY 1968 AAGGATATGAGAGAAACCTGCTTCACTGATGATCTGACAAGGCAACGCGCGAGAGAGTGG 2027
DB 1413 AAGGATATGAGAGAAACCTGCTTCACTGATGATCTGACAAGGCAACGCGCGAGAGAGTGG 1472
QY 2028 AAGTGTGAGGACACACTCTGTGACACAACATCGAGCGGATTTGGCCCTTTACCGAT 2087
DB 1473 AAGTGTGAGGACACACTCTGTGACACAACATCGAGCGGATTTGGCCCTTTACCGAT 1532
QY 2088 GTTCGTTAG 2096
DB 1533 GTTCGTTAG 1541

RESULT 4
US-09-940-235-10
Sequence 10, Application US/09940235
Publication No. US2003065992A1
GENERAL INFORMATION:
APPLICANT: Kumar, Rajesh
APPLICANT: Sahni, Girish
APPLICANT: Roy, Chait
APPLICANT: Rajagopal, Kamnara
APPLICANT: Nihalani, Deepak

APPLICANT: Sundaram, Vasudha
APPLICANT: Yadav, Mahavir
TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
TITLE OF INVENTION: PROTEIN
FILE REFERENCE: 07064-009002
CURRENT APPLICATION NUMBER: US/09/940,235
CURRENT FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: 09/471,349
PRIOR FILING DATE: 1999-12-23
PRIOR APPLICATION NUMBER: IN 3825/DEL/98
PRIOR FILING DATE: 1998-12-24
NUMBER OF SEQ ID NOS: 28
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10
LENGTH: 1661
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Hybrid cassette
US-09-940-235-10

Query Match 55.6%; Score 1165.8; DB 3; Length 1661;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1170; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 588 ATTTCGAGCTGAGGCGCTGAGACCGTCCATCTGTCAACAACGCCAATGTTGTT 647
DB 183 ATTGCTGAGCTGAGGCGCTGAGACCGTCCATCTGTCAACAACGCCAATGTTGTT 242
QY 648 AGCGTGTGCTGAGCTGAGGCGGGAAGATCAAGACATTAGTCTTAAATTTTGAATC 707
DB 243 AGCGTGTGCTGAGCTGAGGCGGGAAGATCAAGACATTAGTCTTAAATTTTGAATC 302
QY 708 GATCTAATCAATCAAGACCTGCTCATGAGAGAAAGACAGAGAGCTTAACTCAAAATCA 767
DB 303 GATCTAATCAATCAAGACCTGCTCATGAGAGAAAGACAGAGAGCTTAACTCAAAATCA 362
QY 768 AATGCTGCTGAGCTGAGGCGGGAAGATCAAGACATTAGTCTTAAATTTTGAATC 827
DB 363 AATGCTGCTGAGCTGAGGCGGGAAGATCAAGACATTAGTCTTAAATTTTGAATC 422
QY 828 AATGCTGCTGAGCTGAGGCGGGAAGATCAAGACATTAGTCTTAAATTTTGAATC 887
DB 423 AATGCTGCTGAGCTGAGGCGGGAAGATCAAGACATTAGTCTTAAATTTTGAATC 482
QY 888 ATTGCTGCTGAGCTGAGGCGGGAAGATCAAGACATTAGTCTTAAATTTTGAATC 947
DB 483 ATTGCTGCTGAGCTGAGGCGGGAAGATCAAGACATTAGTCTTAAATTTTGAATC 542
QY 948 AATGCTGCTGAGCTGAGGCGGGAAGATCAAGACATTAGTCTTAAATTTTGAATC 1007
DB 543 AATGCTGCTGAGCTGAGGCGGGAAGATCAAGACATTAGTCTTAAATTTTGAATC 602
QY 1008 GTCGCGGTGAGCAATATTAAGAAACCAATACAAACCAACCAATATTTTGAATC 1067
DB 603 GTCGCGGTGAGCAATATTAAGAAACCAATACAAACCAACCAATATTTTGAATC 662
QY 1068 GATATCTGCTGAGCTGAGGCGGGAAGATCAAGACATTAGTCTTAAATTTTGAATC 1127
DB 663 GATATCTGCTGAGCTGAGGCGGGAAGATCAAGACATTAGTCTTAAATTTTGAATC 722
QY 1128 GATATCTGCTGAGCTGAGGCGGGAAGATCAAGACATTAGTCTTAAATTTTGAATC 1187
DB 723 GATATCTGCTGAGCTGAGGCGGGAAGATCAAGACATTAGTCTTAAATTTTGAATC 782
QY 1188 GCTCAAGCAAGCAATTTTAAACAAACCAACCAACCAACCAATATTTTGAATC 1247
DB 783 GCTCAAGCAAGCAATTTTAAACAAACCAACCAACCAACCAATATTTTGAATC 842
QY 1248 TCCCTAATGCTGAGCTGAGGCGGGAAGATCAAGACATTAGTCTTAAATTTTGAATC 1307

DB 843 TCCCTAATGCTGAGCTGAGGCGGGAAGATCAAGACATTAGTCTTAAATTTTGAATC 902
QY 1308 TTTACTTACCGTGTAAATATCGGGAACAGCTTATAGATCAATATTAATATCGTGTG 1367
DB 903 TTTACTTACCGTGTAAATATCGGGAACAGCTTATAGATCAATATTAATATCGTGTG 962
QY 1368 AATGAGAAATTAACCAACTGACCTGATCTCTGAGAAATTTTAAATTTTGAATC 1427
DB 963 AATGAGAAATTAACCAACTGACCTGATCTCTGAGAAATTTTAAATTTTGAATC 1022
QY 1428 GAAAAGCGGTATGATCCCTTTGATGCGAGTCACTTGAATCTGTTCAATCAATACGTT 1487
DB 1023 GAAAAGCGGTATGATCCCTTTGATGCGAGTCACTTGAATCTGTTCAATCAATACGTT 1082
QY 1488 GATGCTGATCAACGAATGCTTAAATATGAGAGAGCTTAACTGAGGAGAGCTTAA 1547
DB 1083 GATGCTGATCAACGAATGCTTAAATATGAGAGAGCTTAACTGAGGAGAGCTTAA 1142
QY 1548 TTAGACTTCAAGAGATTTTAAAGATCTGATTAAGCTTAACTTCAACAAATCTC 1607
DB 1143 TTAGACTTCAAGAGATTTTAAAGATCTGATTAAGCTTAACTTCAACAAATCTC 1202
QY 1608 GATGCTTGTGATTAATGAGTATACCTTAACTGAGAAAGTAAAGATTAATCAGATAC 1667
DB 1203 GATGCTTGTGATTAATGAGTATACCTTAACTGAGAAAGTAAAGATTAATCAGATAC 1262
QY 1668 ACCAAGCGGTATCAATACCTTAAATGAGGAGAGAGAGAGAGAGATTAATGATAC 1727
DB 1263 ACCAAGCGGTATCAATACCTTAAATGAGGAGAGAGAGAGAGATTAATGATAC 1322
QY 1728 CATTTAGCTGCTGAGGCGGGAAGATCAAGACATTAGTCTTAAATTTTGAATC 1764
DB 1323 CATTTAGCTGCTGAGGCGGGAAGATCAAGACATTAGTCTTAAATTTTGAATC 1359

RESULT 5

US-09-940-235-5
Sequence 5, Application US/09940235
Publication No. US2003005921A1

GENERAL INFORMATION:
APPLICANT: Kumar, Rajesh
APPLICANT: Sahni, Girish

APPLICANT: Roy, Chait
APPLICANT: Rajagopal, Kammarra

APPLICANT: Nihalani, Deepak
APPLICANT: Sundaram, Vasudha

APPLICANT: Yadav, Mahavir
TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE

TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID

TITLE OF INVENTION: PROTEIN
FILE REFERENCE: 07064-009002

CURRENT APPLICATION NUMBER: US/09/940,235
CURRENT FILING DATE: 2002-04-09

PRIOR APPLICATION NUMBER: 09/471,349
PRIOR FILING DATE: 1999-12-23

PRIOR APPLICATION NUMBER: IN 3825/DEL/98
PRIOR FILING DATE: 1998-12-24

NUMBER OF SEQ ID NOS: 28
SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 5
LENGTH: 1377

TYPE: DNA
ORGANISM: Streptococcus equisimilis

US-09-940-235-5

Query Match 54.7%; Score 1147; DB 3; Length 1377;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1150; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 588 ATTGCTGAGCTGAGGCGCTGAGACCGTCCATCTGTCAACAACGCCAATGTTGTT 647
DB 133 ATTGCTGAGCTGAGGCGCTGAGACCGTCCATCTGTCAACAACGCCAATGTTGTT 192

QY 648 AGCGTTGCTGTAAGTGGGGAAGATCAAGATTAAGTCTTAATTTTGAATC 707
 DB 193 AGCGTTGCTGTAAGTGGGGAAGATCAAGATTAAGTCTTAATTTTGAATC 252
 QY 708 GATCTAACATCAAGCTGCTCATGAGGAAAGACAGAGGCTTAAGTCCAAATCA 767
 DB 253 GATCTAACATCAAGCTGCTCATGAGGAAAGACAGAGGCTTAAGTCCAAATCA 312
 QY 768 AAACCAATTTGCTACTGATAGTGGCGGATGTCATATAAATTGAGAAAGCTGACTTA 827
 DB 313 AAACCAATTTGCTACTGATAGTGGCGGATGTCATATAAATTGAGAAAGCTGACTTA 372
 QY 828 AAGGCTATTCAAGAAATTTGATCGCTAACGTCACAGTAACGACGCTACTTTGAGGTC 887
 DB 373 AAGGCTATTCAAGAAATTTGATCGCTAACGTCACAGTAACGACGCTACTTTGAGGTC 432
 QY 888 ATTGATTTTGAAGCGATGCAACCATTAAGTGAAGCGGCAAGGCTTAAGTGGCTGAC 947
 DB 433 ATTGATTTTGAAGCGATGCAACCATTAAGTGAAGCGGCAAGGCTTAAGTGGCTGAC 492
 QY 948 AAAGATGCTTGGTAACTTTGCCGACCCAACTGTCAGAAATTTTGTCTAAGCGGACAT 1007
 DB 493 AAAGATGCTTGGTAACTTTGCCGACCCAACTGTCAGAAATTTTGTCTAAGCGGACAT 552
 QY 1008 GTGGCGCTTGAAGCAATATAAGAAACCAATACAAACGCAAGAAATGCTGATGTC 1067
 DB 553 GTGGCGCTTGAAGCAATATAAGAAACCAATACAAACGCAAGAAATGCTGATGTC 612
 QY 1068 GAATATCTGTAAGTCTTCTCCCTTAAACCTGATGAGATTTTCAGACCAAGGCTCAAA 1127
 DB 613 GAATATCTGTAAGTCTTCTCCCTTAAACCTGATGAGATTTTCAGACCAAGGCTCAAA 672
 QY 1128 GATCTAACATCAAGCTGCTCATGAGGAAAGACAGATTAAGTCTTAATTTTGAATC 1187
 DB 673 GATCTAACATCAAGCTGCTCATGAGGAAAGACAGATTAAGTCTTAATTTTGAATC 732
 QY 1188 GCTCAAGCAAGAAATTTTAAACCAACCCGCGCTATACGATTAAGTGAAGCGGAC 1247
 DB 733 GCTCAAGCAAGAAATTTTAAACCAACCCGCGCTATACGATTAAGTGAAGCGGAC 792
 QY 1248 TCCCTAATGCTCATGATGACATGATCTTCCGATGATTTTCAATGATGATCAAG 1307
 DB 793 TCCCTAATGCTCATGATGACATGATCTTCCGATGATTTTCAATGATGATCAAG 852
 QY 1308 TTTACTTACCGTGTAAATTCGGGAAACAAGCTTAATGATCAATTAATTTTGAATC 1367
 DB 853 TTTACTTACCGTGTAAATTCGGGAAACAAGCTTAATGATCAATTAATTTTGAATC 912
 QY 1368 AATGAAGAAATTAACCAAGCTGATCTGAGAAATTTTACGCTTTAAAGG 1427
 DB 913 AATGAAGAAATTAACCAAGCTGATCTGAGAAATTTTACGCTTTAAAGG 972
 QY 1428 GAAAGCGGTATGATCCCTTGAAGCAGTCACTTGAACCTGTTCAATCAATGATG 1487
 DB 973 GAAAGCGGTATGATCCCTTGAAGCAGTCACTTGAACCTGTTCAATCAATGATG 1032
 QY 1488 GATGTCGATCAAGCAATTTGCTAAAGAGTGAAGCAGCTTTTACAGCTAGCGAAG 1547
 DB 1033 GATGTCGATCAAGCAATTTGCTAAAGAGTGAAGCAGCTTTTACAGCTAGCGAAG 1092
 QY 1548 TTAAGCTTGAAGATTTTATGAGATCTGCTGATAGGCTTAAGTCTTAAGTCTGAC 1607
 DB 1093 TTAAGCTTGAAGATTTTATGAGATCTGCTGATAGGCTTAAGTCTTAAGTCTGAC 1152
 QY 1608 GATGCTTTGATTAATGAGATTAATGAGTGAAGGATTAAGTGAAGTGAAGTGAAG 1667
 DB 1153 GATGCTTTGATTAATGAGATTAATGAGTGAAGGATTAAGTGAAGTGAAGTGAAG 1212
 QY 1668 ACCAAGCGGTATGATCAAGCTTTTATGAGGCAAGCGAAGCGAAGGATGCTAGTAC 1727
 DB 1213 ACCAAGCGGTATGATCAAGCTTTTATGAGGCAAGCGAAGCGAAGGATGCTAGTAC 1272

QY 1728 CATTACCTGATGCT 1742
 DB 1273 CATTACCTGATGCT 1287
 RESULT 6
 US-09-940-235-1
 ; Sequence 1, Application US/09940235
 ; Publication No. US2003005921A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kumar, Rajesh
 ; APPLICANT: Sahni, Girdish
 ; APPLICANT: Roy, Chait
 ; APPLICANT: Rajagopal, Kammaru
 ; APPLICANT: Nihalani, Deepak
 ; APPLICANT: Sundaram, Vasudha
 ; APPLICANT: Yadav, Mahavir
 ; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
 ; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
 ; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
 ; TITLE OF INVENTION: PROTEIN
 ; FILE REFERENCE: 07064-009002
 ; CURRENT APPLICATION NUMBER: US/09/940,235
 ; CURRENT FILING DATE: 2002-04-09
 ; PRIOR APPLICATION NUMBER: 09/471,349
 ; PRIOR FILING DATE: 1999-12-23
 ; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
 ; PRIOR FILING DATE: 1998-12-24
 ; NUMBER OF SEQ ID NOS: 28
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 1245
 ; TYPE: DNA
 ; ORGANISM: Streptococcus equisimilis
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)...(1242)
 US-09-940-235-1
 Query Match 54.6%; Score 1145.4; DB 3; Length 1245;
 Best Local Similarity 99.5%; Pred. No. 0;
 Matches 1149; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 588 ATTGCTGACCTGATAGTGGCTGCTAGACCGTCCATCTGTCAACACAGGCAATGCTTGT 647
 DB 1 ATTGCTGACCTGATAGTGGCTGCTAGACCGTCCATCTGTCAACACAGGCAATGCTTGT 60
 QY 648 AGCGTTGCTGTAAGTGGGGAAGATCAAGATTAAGTCTTAATTTTGAATC 707
 DB 61 AGCGTTGCTGTAAGTGGGGAAGATCAAGATTAAGTCTTAATTTTGAATC 120
 QY 708 GATCTAACATCAAGCTGCTCATGAGGAAAGACAGAGGCTTAAGTCCAAATCA 767
 DB 121 GATCTAACATCAAGCTGCTCATGAGGAAAGACAGAGGCTTAAGTCCAAATCA 180
 QY 768 AAACCAATTTGCTACTGATAGTGGCGGATGTCATATAAATTGAGAAAGCTGACTTA 827
 DB 181 AAACCAATTTGCTACTGATAGTGGCGGATGTCATATAAATTGAGAAAGCTGACTTA 240
 QY 828 AAGGCTATTCAAGAAATTTGATCGCTAACGTCACAGTAACGACGCTACTTTGAGGTC 887
 DB 241 AAGGCTATTCAAGAAATTTGATCGCTAACGTCACAGTAACGACGCTACTTTGAGGTC 300
 QY 888 ATTGATTTTGAAGCGATGCAACCATTAAGTGAAGCGGCAAGGCTTAAGTGGCTGAC 947
 DB 301 ATTGATTTTGAAGCGATGCAACCATTAAGTGAAGCGGCAAGGCTTAAGTGGCTGAC 360
 QY 948 AAAGATGCTTGGTAACTTTGCCGACCCAACTGTCAGAAATTTTGTCTAAGCGGACAT 1007
 DB 361 AAAGATGCTTGGTAACTTTGCCGACCCAACTGTCAGAAATTTTGTCTAAGCGGACAT 420
 QY 1008 GTGGCGCTTGAAGCAATATAAGAAACCAATACAAACGCAAGAAATGCTGATGTC 1067


```

Qy 1488 GATGTCGATACCAACGAATTTGCTAAAAAGTACAGCTCTTAAACAGCTACGAAACGTTAC 1547
Db 983 GATGTCGATACCAACGAATTTGCTAAAAAGTACAGCTCTTAAACAGCTACGAAACGTTAC 1042
Qy 1548 TTAGACTTCAAGATTTTAAACGATCTCTCGTATTAAGGCTTAACTACTTCAACATCTC 1607
Db 1043 TTAGACTTCAAGATTTTAAACGATCTCTCGTATTAAGGCTTAACTACTTCAACATCTC 1102
Qy 1608 GATGCTTTGGTATTATTAAGCTATTAACCTTAACTGAAAGTAAAGTAAATTCACGATAC 1667
Db 1103 GATGCTTTGGTATTATTAAGCTATTAACCTTAACTGAAAGTAAAGTAAATTCACGATAC 1162
Qy 1668 ACCAACCGTATATTAACCGTTTATATGCGACAGCCGCAAGAGAGAAATGCTAGCTAC 1727
Db 1163 ACCAACCGTATATTAACCGTTTATATGCGACAGCCGCAAGAGAGAAATGCTAGCTAT 1222
Qy 1728 CATTTAGCTGTGTGT 1742
Db 1223 CATTTAGCTGTGTAT 1237

```

RESULT 8

```

US-10-474-792-657
; Sequence 657, Application US/10474792
; Publication No. US20040236072A1
; GENERAL INFORMATION:
; APPLICANT: Olmsted, Stephen
; APPLICANT: Zagursky, Robert
; APPLICANT: Nickbarg, Elliot
; APPLICANT: Winer, Laurie
; TITLE OF INVENTION: SURFACE PROTEINS OF STREPTOCOCCUS PYOGENES
; FILE REFERENCE: AM 100399
; CURRENT APPLICATION NUMBER: US/10/474,792
; CURRENT FILING DATE: 2003-10-14
; NUMBER OF SEQ ID NOS: 674
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 657
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Streptococcus pyogenes
US-10-474-792-657

```

```

Query Match      47.5%; Score 996; DB 8; Length 1323;
Best Local Similarity 91.3%; Pred. No. 1.9e-277;
Matches 1056; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

```

```

Qy 587 TATGCTGAGCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAGCCAAATTTGGTGT 646
Db 78 TATGCTGAGCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAGCCAAATTTGGTGT 137
Qy 647 TAGCGTTGCTGTAAGTGTGAGGGGAGCAATCAAGCATTAAGTCTTAAATTTTGAAT 706
Db 138 TAGTATGCGCGGTATCGTTGAAGGTACCGATTAAGTTTAAATTTTGAAT 197
Qy 707 CGATCTAATCATGACGCTGCTCATGAGGAAAGAGAGGCAAGCTTAAATCCAAATC 766
Db 198 CGATCTAATCATGACGCTGCTCATGAGGAAAGAGAGGCAAGCTTAAATCCAAATC 257
Qy 767 AAAACCATTTGCTACTGATAGTGGCGCATGTCAATTAACCTTGAAGAAAGCTGACTTACT 826
Db 258 AAAACCATTTGCTACTGATAGTGGCGCATGTCAATTAACCTTGAAGAAAGCTGACTTACT 317
Qy 827 AAAAGCTATTCAGAAACAATTAATGCTTAACGTCACAGTAACGACGACTCTTTGAGGT 886
Db 318 AAAAGCTATTCAGAAACAATTAATGCTTAACGTCACAGTAACGACGACTCTTTGAGGT 377
Qy 887 CATTTATTTTGAAGAGGCAAGCAATTAATGCTTAACGTCACAGTAACGACGACTCTTTGAGGT 946
Db 378 CATTTATTTTGAAGAGGCAAGCAATTAATGCTTAACGTCACAGTAACGACGACTCTTTGAGGT 437
Qy 947 CAAGAATGTTGCGTAACCTTCCGACCAACCTGTCACAAATTTTGTAGACGACA 1006
Db 438 CAAGAATGTTGCGTAACCTTCCGACCAACCTGTCACAAATTTTGTAGACGACA 497

```

```

Qy 1007 TGTGCGGTTAGACCATATTAAGAAAAACCAATATCAAAACCAAGCAATCTGTTGATGT 1066
Db 498 TGTGCGGTTAGACCATATTAAGAAAAACCAATATCAAAACCAAGCAATCTGTTGATGT 557
Qy 1067 GGAATATATCTGTACAGTTTACTCCCTTAAACCTGATGAGATTTGACAGGCTTCAA 1126
Db 558 AGAATATATCTGTACAGTTTACTCCCTTAAACCTGATGAGATTTGACAGGCTTCAA 617
Qy 1127 AGATCTAATCTATTTGAAAAACCTAGTATCGGTGACACATCACTCATCTCAAGATTAAT 1186
Db 618 AGATCTAATCTATTTGAAAAACCTAGTATCGGTGACACATCACTCATCTCAAGATTAAT 677
Qy 1187 AGCTCAAGCAACAAAGCTTTTAAACAAACCCAGGCTATATGATTAAGAACGTGA 1246
Db 678 AGCTCAAGCAACAAAGCTTTTAAACAAACCCAGGCTATATGATTAAGAACGTGA 737
Qy 1247 CTCCTCAATGCTGACCTGATGACATGACATTTTCCGTACGATTTTACCAATGATCAAGA 1306
Db 738 CTCCTCAATGCTGACCTGATGACATGACATTTTCCGTACGATTTTACCAATGATCAAGA 797
Qy 1307 GTTACTTACCGTGTAAAAATGGGAAACAGCTTATAGATCAATTAATAAATCTGCT 1366
Db 798 GTTACTTACCGTGTAAAAATGGGAAACAGCTTATAGATCAATCTTAACAGGTAT 857
Qy 1367 GAATGAAGAATTAACAACACGACCTGATCTGGAATTAATTAAGCTCTTAAAAAAG 1426
Db 858 TAAAGAAAAACCAACCAACCTGATCTGGAATTAATTAAGCTCTTAAAAAAG 917
Qy 1427 GAAAAAGCCGTATGATCCCTTGTATGTCAGCTCACTTGAACCTGTTCAACCAATATAGT 1486
Db 918 GAAAAAGCCGTATGATCCCTTGTATGTCAGCTCACTTGAACCTGTTCAACCAATATAGT 977
Qy 1487 TGATGTCATCAACCAACCAATGCTTAAAAAGTGAAGAGCTTTTAAACAGTACGAACTTAA 1546
Db 978 TGATGTCATCAACCAACCAATGCTTAAAAAGTGAAGAGCTTTTAAACAGTACGAACTTAA 1037
Qy 1547 CTTAAGCTTCAAGATTTTAAACGATCTCTGATTAAGGCTTAAACTACTTCAACCAATCT 1606
Db 1038 CTTAAGCTTCAAGATTTTAAACGATCTCTGATTAAGGCTTAAACTACTTCAACCAATCT 1097
Qy 1607 CGATGCTTTGATATTAATGACCTTAACTGAAAAAGTGAAGATATATCAAGTAA 1666
Db 1098 CGATGCTTTGATATTAATGACCTTAACTGAAAAAGTGAAGATATATCAAGTAA 1157
Qy 1667 CACCAACCGTATATTAACCGTTTATATGAGCAAGCCGCAAGAGAGAGATGCTAGCTA 1726
Db 1158 GAATTAATCGTGTGTTAATGATTAATGAGCAAGCCGCTTAAAGGGCAAGGCTAGCTA 1217
Qy 1727 CCATTTAGCTGTGTGT 1742
Db 1218 TCATTTAGCTTATGAT 1233

```

RESULT 9

```

US-09-940-235-3
; Sequence 3, Application US/09940235
; Publication No. US20030059921A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Gajish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammarra
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Vadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/09/940,235
; CURRENT FILING DATE: 2002-04-09

```


RESULT 15
US-10-741-601-75
; Sequence 75, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001500
; CURRENT APPLICATION NUMBER: US/10/741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 75
; LENGTH: 2488
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-601-75

Query Match 15.7%; Score 330; DB 7; Length 2488;
Best Local Similarity 100.0%; Pred. No. 4.3e-84;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1764 CCCATGCTGAGAGTGTGTTGATCATGCTGCGACTTCTATGTGTGCGAGAAACG 1823
|||
DB 913 CCCATGCTGAGAGTGTGTTGATCATGCTGCGACTTCTATGTGTGCGAGAAACG 972
|||
QY 1824 TGGGAGAAAGCCCTACCAAGCTGATGATGTAGATTGTACTTGCTGGAGAAAGGCAAGC 1883
|||
DB 973 TGGGAGAAAGCCCTACCAAGCTGATGATGTAGATTGTACTTGCTGGAGAAAGGCAAGC 1032
|||
QY 1884 GGACGATCACTTGACCTTCTAGAAATAGATGCAACGATCAGACACAAAGACATCCTAT 1943
|||
DB 1033 GGACGATCACTTGACCTTCTAGAAATAGATGCAACGATCAGACACAAAGACATCCTAT 1092
|||
QY 1944 AGAATTGAGACACCTGAGCAAGAGATATCGAGAAACCTGCTCCAGTGCATCTGC 2003
|||
DB 1093 AGAATTGAGACACCTGAGCAAGAGATATCGAGAAACCTGCTCCAGTGCATCTGC 1152
|||
QY 2004 ACAAGCAAGCCCGAGAGAGAGTGAAGTGTGAGAGGCAACCTCTGTGCAACCAATCG 2063
|||
DB 1153 ACAAGCAAGCCCGAGAGAGAGTGAAGTGTGAGAGGCAACCTCTGTGCAACCAATCG 1212
|||
QY 2064 AGCGGATCTGGCCCTTCAACCGATGTTGGT 2093
|||
DB 1213 AGCGGATCTGGCCCTTCAACCGATGTTGGT 1242
|||

Search completed: February 1, 2006, 14:21:22
Job time : 1732.86 secs

Db 61 TTCGCTCGGTAATCGGTGATTAATCTGCTAACAGTAAGCAACCCCGCAGCTAGCC 120
 Qy 121 GGGTCTCAACGACAGAGACAGATCATGCGACCCGTGGCCAGAACCCAAAGCTGGCCG 180
 Db 121 GGGTCTCAACGACAGAGACAGATCATGCGACCCGTGGCCAGAACCCAAAGCTGGCCG 180
 Qy 181 AGATCTCGATCCCGGCAAAATTAATACGACTCACTATAGGGAGACCAACGCTTTCCCTC 240
 Db 181 AGATCTCGATCCCGGCAAAATTAATACGACTCACTATAGGGAGACCAACGCTTTCCCTC 240
 Qy 241 TAGAAATATTTTGTAACTTTAAAGAGAGATATACCTATGTCAGAGACCAACAGATT 300
 Db 241 TAGAAATATTTTGTAACTTTAAAGAGAGATATACCTATGTCAGAGACCAACAGATT 300
 Qy 301 GTACCACTAGCTGAGAAAGTGTATGATCATGCTGAGACTTCTATGTCGAGAA 360
 Db 301 GTACCACTAGCTGAGAAAGTGTATGATCATGCTGAGACTTCTATGTCGAGAA 360
 Qy 361 ACGTGGAGAGAGGACGAGACGATCATCTTCTAGAAATAGATGCAACGATCAGG 420
 Db 361 ACGTGGAGAGAGGACGAGACGATCATCTTCTAGAAATAGATGCAACGATCAGG 420
 Qy 421 ACAGAGAGACATCTTATAGAAATTTGAGACACCTGAGCAAGAGATATGAGAAACC 480
 Db 421 ACAGAGAGACATCTTATAGAAATTTGAGACACCTGAGCAAGAGATATGAGAAACC 480
 Qy 481 TGCTCAAGTCACTGTCACAGGCAAGCGAGAGAGTGGAGAGGCAACGACT 540
 Db 481 TGCTCAAGTCACTGTCACAGGCAAGCGAGAGAGTGGAGAGGCAACGACT 540
 Qy 541 CTGTGCAAGACCATGAGCGGATCTGGCCCTTCAACCATGTTGCTGAGACCTG 600
 Db 541 CTGTGCAAGACCATGAGCGGATCTGGCCCTTCAACCATGTTGCTGAGACCTG 600
 Qy 601 AGTGGCTGTAAACGCTCATCTGTCAACAACAGCCAATTTGTTAGGCTTGGTGA 660
 Db 601 AGTGGCTGTAAACGCTCATCTGTCAACAACAGCCAATTTGTTAGGCTTGGTGA 660
 Qy 661 CTGTTAGGGGAGCAATCAAGACATTAAGTCTTAAATTTTGAATCGATCTAACATC 720
 Db 661 CTGTTAGGGGAGCAATCAAGACATTAAGTCTTAAATTTTGAATCGATCTAACATC 720
 Qy 721 GACCTGCTCATGAGAGAAAGACAGAGCAAGCTTAAATCAAAATCAAAATCTTGTCTA 780
 Db 721 GACCTGCTCATGAGAGAAAGACAGAGCAAGCTTAAATCAAAATCAAAATCTTGTCTA 780
 Qy 781 CTGATAGTGGCGGATGTCAATTAACCTTGAAGAGCTGAATTAAAGGCTATTCAG 840
 Db 781 CTGATAGTGGCGGATGTCAATTAACCTTGAAGAGCTGAATTAAAGGCTATTCAG 840
 Qy 841 AACATTTGATGCTTAACGCTTCAACAGTAAGCACTTATGAGTCAATTTTGCAG 900
 Db 841 AACATTTGATGCTTAACGCTTCAACAGTAAGCACTTATGAGTCAATTTTGCAG 900
 Qy 901 GCGATGCAACCATTAATGATCGAAACGCAAGGTCTACTTGTCTGCAAAAGTGTTCG 960
 Db 901 GCGATGCAACCATTAATGATCGAAACGCAAGGTCTACTTGTCTGCAAAAGTGTTCG 960
 Qy 961 TAACTTGTGCGAACCAACCTGTGCAAGAAATTTTGTGAAGCGGACATGTGCGTTAG 1020
 Db 961 TAACTTGTGCGAACCAACCTGTGCAAGAAATTTTGTGAAGCGGACATGTGCGTTAG 1020
 Qy 1021 CATATTAAGAAAAACCAATACAAACCAAGCGAAATCTGTGATGTGGAATATATCTGAC 1080
 Db 1021 CATATTAAGAAAAACCAATACAAACCAAGCGAAATCTGTGATGTGGAATATATCTGAC 1080
 Qy 1081 AGTTACTCCCTTAAACCTGTGATGAGATTTGAGACAGGCTCAAGATCTAAGCTAT 1140
 Db 1081 AGTTACTCCCTTAAACCTGTGATGAGATTTGAGACAGGCTCAAGATCTAAGCTAT 1140
 Qy 1141 TGAAGAACATAGCTATCGTGAACCATCATCTCAAGAAATTAAGTCTAGGCAACAA 1200
 Db 1141 TGAAGAACATAGCTATCGTGAACCATCATCTCAAGAAATTAAGTCTAGGCAACAA 1200

Qy 1201 GCATTTTAAACMAAAACCAACGCTATACGATTTATGAAGTGAATCTCAATCTCA 1260
 Db 1201 GCATTTTAAACMAAAACCAACGCTATACGATTTATGAAGTGAATCTCAATCTCA 1260
 Qy 1261 CTATGACATGACATTTTCCGTACGATTTTACCAATGATCAAGAGTTTACTTACG 1320
 Db 1261 CTATGACATGACATTTTCCGTACGATTTTACCAATGATCAAGAGTTTACTTACG 1320
 Qy 1321 TTAATAATCGGGAACAAGCTTATAGGATCAATTAATAATCTGCTGAATGAAGAAATA 1380
 Db 1321 TTAATAATCGGGAACAAGCTTATAGGATCAATTAATAATCTGCTGAATGAAGAAATA 1380
 Qy 1381 ACAACACTGACCTGATCTGAGAAATTAATGCTCTTAAATAAGGGGAAAAAGCGTATG 1440
 Db 1381 ACAACACTGACCTGATCTGAGAAATTAATGCTCTTAAATAAGGGGAAAAAGCGTATG 1440
 Qy 1441 ATCCCTTGAATCGAGTCACTTGAACCTGTTCAACATCAAAATAGCTGATGATACCA 1500
 Db 1441 ATCCCTTGAATCGAGTCACTTGAACCTGTTCAACATCAAAATAGCTGATGATACCA 1500
 Qy 1501 ACGAATGCTTAAAGTGAAGAGCTCTTAAACAGCTAGGAACTTAACTTAACTTCAAG 1560
 Db 1501 ACGAATGCTTAAAGTGAAGAGCTCTTAAACAGCTAGGAACTTAACTTAACTTCAAG 1560
 Qy 1561 ATTTATAGATCCCTGATTAAGGCTTAACTTAACTTCAACATCTGATCTTGGTGA 1620
 Db 1561 ATTTATAGATCCCTGATTAAGGCTTAACTTAACTTCAACATCTGATCTTGGTGA 1620
 Qy 1621 TTATGACATTAACCTTAACTGAGAAAGTGAAGATTAATCAAGATGACCAACCGTATCA 1680
 Db 1621 TTATGACATTAACCTTAACTGAGAAAGTGAAGATTAATCAAGATGACCAACCGTATCA 1680
 Qy 1681 TTAACGTTTATATGGGCAAGGACCCGAGAGAGATGCTAGCTACATTTAGTGGTG 1740
 Db 1681 TTAACGTTTATATGGGCAAGGACCCGAGAGAGATGCTAGCTACATTTAGTGGTG 1740
 Qy 1741 GTGGCAGAGGCAACAGATGTATACCAATGCTGAAGAGTGTGTTGATCATGCTGCGGA 1800
 Db 1741 GTGGCAGAGGCAACAGATGTATACCAATGCTGAAGAGTGTGTTGATCATGCTGCGGA 1800
 Qy 1801 CTTCCTATGTTGTTGCGAGAAACGTGGAGAAAGCCCTTCAAGGCTGATGATGTAAGTT 1860
 Db 1801 CTTCCTATGTTGTTGCGAGAAACGTGGAGAAAGCCCTTCAAGGCTGATGATGTAAGTT 1860
 Qy 1861 GTACTTGCCTGGAGAGAGGACGCGACATCACTTGCATTTCTAGAAATGATGCAACG 1920
 Db 1861 GTACTTGCCTGGAGAGAGGACGCGACATCACTTGCATTTCTAGAAATGATGCAACG 1920
 Qy 1921 ATCAGGACACAGGACATCTATGAAATTTGAGACACCTGAGCAAGAGATTAATGAG 1980
 Db 1921 ATCAGGACACAGGACATCTATGAAATTTGAGACACCTGAGCAAGAGATTAATGAG 1980
 Qy 1981 GAAACCTGCTCAGTGAATCTGACAGGCAACGCGGAGAGAGTGAAGTGTGAGAGGC 2040
 Db 1981 GAAACCTGCTCAGTGAATCTGACAGGCAACGCGGAGAGAGTGAAGTGTGAGAGGC 2040
 Qy 2041 ACACCTTGTGACAGACCATGAGCGGATCTGACCCCTTCAACGATGTTGTTAG 2096
 Db 2041 ACACCTTGTGACAGACCATGAGCGGATCTGACCCCTTCAACGATGTTGTTAG 2096

RESULT 2

US-10-631-558-11
 ; Sequence 11, Application US/10631558
 ; Publication No. US20050260598A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kumar, Rajesh
 ; APPLICANT: Sami, Girish
 ; APPLICANT: Roy, Chait
 ; APPLICANT: Rajagopal, Kammarra
 ; APPLICANT: Nihalani, Deepak
 ; APPLICANT: Sundaram, Vasudha

APPLICANT: Yadav, Mahavir
TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
FILE OF INVENTION: PROTEIN
FILE REFERENCE: 07064-009002
CURRENT APPLICATION NUMBER: US/10/631,558
CURRENT FILING DATE: 2003-07-31
PRIOR APPLICATION NUMBER: US/09/940,235
PRIOR FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: 09/471,349
PRIOR FILING DATE: 1999-12-23
PRIOR APPLICATION NUMBER: IN 3825/DEL/98
PRIOR FILING DATE: 1998-12-24
NUMBER OF SEQ ID NOS: 28
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 11
LENGTH: 1782
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Hybrid cassette
US-10-631-558-11

Query Match 80.3%; Score 1684; DB 7; Length 1782;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1687; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 51 TCGCTTACGTTCCGCTCGGCTATCGGTATTCATTCTGCTAACCAAGTAAAGCAACCCCGC 110
DB 1 TCGCTTACGTTCCGCTCGGCTATCGGTATTCATTCTGCTAACCAAGTAAAGCAACCCCGC 60

QY 111 CAGCTAGCCGGGCTCTCAACGACAGAGACAGATCATGCGCACCCGTGCGCAGACCA 170
DB 61 CAGCTAGCCGGGCTCTCAACGACAGAGACAGATCATGCGCACCCGTGCGCAGACCA 120

QY 171 ACGCTGCCGAGATCTCGATCCCGCGAAATTAATAGACTCATATAGGAGACCAAC 230
DB 121 ACGCTGCCGAGATCTCGATCCCGCGAAATTAATAGACTCATATAGGAGACCAAC 180

QY 231 GGTTCCTCTAGAAATTAATTTTGTATTAATTAAAGAGATATACATGCTGCAAGC 290
DB 181 GGTTCCTCTAGAAATTAATTTTGTATTAATTAAAGAGATATACATGCTGCTGATGT 240

QY 291 ACAAGATATGTAACCATAGCTGAGAAAGTTTGAATCATGCTGCGGACTTCTATGT 350
DB 241 ACAAGATATGTAACCATAGCTGAGAAAGTTTGAATCATGCTGCGGACTTCTATGT 300

QY 351 GGTGAGAGAAAGTGGAGAGAGGAGCGACGATCATCTTGACCTTGAATATGATGC 410
DB 301 GGTGAGAGAAAGTGGAGAGAGGAGCGACGATCATCTTGACCTTGAATATGATGC 360

QY 411 AACGATCAGACACAAAGACATCTTATAGATTTGAGACACTGAGACAAAGATTAAT 470
DB 361 AACGATCAGACACAAAGACATCTTATAGATTTGAGACACTGAGACAAAGATTAAT 420

QY 471 CGAGGAAAACCTGCTCAAGTGCATCTGCAACGACGCGCGAGAGAGTGAAGTGTAG 530
DB 421 CGAGGAAAACCTGCTCAAGTGCATCTGCAACGCGCGAGAGAGTGAAGTGTAG 480

QY 531 AGGCAACCTCTGTCAGACCAATGACCGGATCTGGCCCTTCAACCGATTTGCTATTT 590
DB 481 AGGCAACCTCTGTCAGACCAATGACCGGATCTGGCCCTTCAACCGATTTGCTATTT 540

QY 591 GCTGACCTGAGTGGCTGCTAGACCGTCATCTGCAACAGCCAAATTTGTTTAC 650
DB 541 GCTGACCTGAGTGGCTGCTAGACCGTCATCTGCAACAGCCAAATTTGTTTAC 600

QY 651 GTTGTGTAAGTGTGAAGGAGCAATCAAGACATTAGTCTTAAATTTTGAATCAT 710
DB 601 GTTGTGTAAGTGTGAAGGAGCAATCAAGACATTAGTCTTAAATTTTGAATCAT 660

QY 711 CTAACATCAACGCTGCTCATGAGAGAAAGACAGACAAAGGCTTAAGTCCAAATCAAAA 770

DB 661 CTAACATCAACGCTGCTCATGAGAGAAAGACAGACAAAGGCTTAAGTCCAAATCAAAA 720
QY 771 CCATTTGTAATGTAAGTGGGCGCATGCAATAACTTGAAGAGCTGACTTAAAG 830
DB 721 CCATTTGTAATGTAAGTGGGCGCATGCAATAACTTGAAGAGCTGACTTAAAG 780

QY 831 GCTATTCAGAAACATTAAGTGTGCTTACAGTAAAGAGCTGACTTGAAGTCAAT 890
DB 781 GCTATTCAGAAACATTAAGTGTGCTTACAGTAAAGAGCTGACTTGAAGTCAAT 840

QY 891 GATTTTCAGACGATGCAACCATTAAGTATGAAACGCGCAAGTCTATTGCTGACAAA 950
DB 841 GATTTTCAGACGATGCAACCATTAAGTATGAAACGCGCAAGTCTATTGCTGACAAA 900

QY 951 GATGCTCGTACCTTGGCCGACCACTGCTCAAGAAATTTTCTAAGCGGACATGTG 1010
DB 901 GATGCTCGTACCTTGGCCGACCACTGCTCAAGAAATTTTCTAAGCGGACATGTG 960

QY 1011 CGGTTAGACCATTAAGAAAGAAACCAATCAAAACCAAGGGAATCTGTGATGTGAA 1070
DB 961 CGGTTAGACCATTAAGAAAGAAACCAATCAAAACCAAGGGAATCTGTGATGTGAA 1020

QY 1071 TATACGTACAGTTTACTCCCTTAAACCTGATGACGATTCAGACAGTCTCAAAAGAT 1130
DB 1021 TATACGTACAGTTTACTCCCTTAAACCTGATGACGATTCAGACAGTCTCAAAAGAT 1080

QY 1131 ACTAAGCTATTGAAACATAGCTATGCTGACACCATCACTCAAGAAATTAAGTCT 1190
DB 1081 ACTAAGCTATTGAAACATAGCTATGCTGACACCATCACTCAAGAAATTAAGTCT 1140

QY 1191 CAAGCAAGACATTTTAAACAAACCAACCAAGGCTATAGATTTAAGACGTGATCC 1250
DB 1141 CAAGCAAGACATTTTAAACAAACCAACCAAGGCTATAGATTTAAGACGTGATCC 1200

QY 1251 TCAATCGTCACTGACATGACATTAATTTCCGTCAGATTTTAAACCAAGTCAAGATTT 1310
DB 1201 TCAATCGTCACTGACATGACATTAATTTCCGTCAGATTTTAAACCAAGTCAAGATTT 1260

QY 1311 ACTTACGCTGTTAAATTCGGGAAACAGCTTATAGATCAATTAATAATCTGCTGAT 1370
DB 1261 ACTTACGCTGTTAAATTCGGGAAACAGCTTATAGATCAATTAATAATCTGCTGAT 1320

QY 1371 GAAGAAATTAACCAACATGACCTGATCTGAGAAATTTACGTCCTTAAAAAGGGAA 1430
DB 1321 GAAGAAATTAACCAACATGACCTGATCTGAGAAATTTACGTCCTTAAAAAGGGAA 1380

QY 1431 AAGCCGTAATGATCCCTTGAATGCAATCACTGAAACGTTCAACATTAAGTGTAT 1490
DB 1381 AAGCCGTAATGATCCCTTGAATGCAATCACTGAAACGTTCAACATTAAGTGTAT 1440

QY 1491 GTGATACCAAGAAATGCTTAAAGAGAGAGCTTAAACAGCTAGGAAAGTAACTTA 1550
DB 1441 GTGATACCAAGAAATGCTTAAAGAGAGAGCTTAAACAGCTAGGAAAGTAACTTA 1500

QY 1551 GACTTCAGAGATTTATAGATCTGTCGATTAAGGCTTAACTTCAACAAATCTGAT 1610
DB 1501 GACTTCAGAGATTTATAGATCTGTCGATTAAGGCTTAACTTCAACAAATCTGAT 1560

QY 1611 GCTTTGTAATTAAGCTATTAACCTTAACTGAAAGAGTAAAGATTAACAGATGACACC 1670
DB 1561 GCTTTGTAATTAAGCTATTAACCTTAACTGAAAGAGTAAAGATTAACAGATGACACC 1620

QY 1671 AACCGTATCAATACCGTTTATATAGGAGAGGACCGGAAAGAGAAATGCTAGTACAT 1730
DB 1621 AACCGTATCAATACCGTTTATATAGGAGAGGACCGGAAAGAGAAATGCTAGTACAT 1680

QY 1731 TTAGCTGTGAT 1742
DB 1681 TTAGCTGTGAT 1692

RESULT 3

TYPE: DNA
ORGANISM: Streptococcus equisimilis
US-10-631-558-5

Query Match 54.7%; Score 1147; DB 7; Length 1377;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1150; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

QY 588 ATTGCTGAGCCTGAGGCGTCTAGACCGTCCATCTGTGACAAACAGCCAAATGGTGT 647
DB 133 ATTGCTGAGCCTGAGGCGTCTAGACCGTCCATCTGTGACAAACAGCCAAATGGTGT 192
QY 648 AGCGTTGCTGCTACTGTTGAGGGGAGCAATCAAGACTTATGTTAAATTTTGAATC 707
DB 193 AGCGTTGCTGCTACTGTTGAGGGGAGCAATCAAGACTTATGTTAAATTTTGAATC 252
QY 708 GATCTAACATCAAGACCTGCTCATGAGGAGAAAGACAGCAAGCTTAAAGTCCAAATCA 767
DB 253 GATCTAACATCAAGACCTGCTCATGAGGAGAAAGACAGCAAGCTTAAAGTCCAAATCA 312
QY 768 AAACCAATTTGCTACTGATAGTGGCGGATGTCAATAACTTGAGAAAGTGACTTACTA 827
DB 313 AAACCAATTTGCTACTGATAGTGGCGGATGTCAATAACTTGAGAAAGTGACTTACTA 372
QY 828 AAGGCTATTCAAGAACAAATTTGATCGCTAACGTCCACAGTAAAGCACTACTTGAAGTC 887
DB 373 AAGGCTATTCAAGAACAAATTTGATCGCTAACGTCCACAGTAAAGCACTACTTGAAGTC 432
QY 888 ATTGATTTTGCAGGCGATGCAACCATTAAGTATGAAACGGCAAGGTCTACTTGGCTGAC 947
DB 433 ATTGATTTTGCAGGCGATGCAACCATTAAGTATGAAACGGCAAGGTCTACTTGGCTGAC 492
QY 948 AAAGATGTTGGCTGTAACCTTGGCGACCCACCTGTCCAGAAATTTTGGTGAAGCGACAT 1007
DB 493 AAAGATGTTGGCTGTAACCTTGGCGACCCACCTGTCCAGAAATTTTGGTGAAGCGACAT 552
QY 1008 GTGCGCGTTAGACCATATTAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTC 1067
DB 553 GTGCGCGTTAGACCATATTAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTC 612
QY 1068 GAATTTACTGTAAGTTCCTGCTTAAACCCCTGATGACGATTTTGACAGCAGGTCTCAA 1127
DB 613 GAATTTACTGTAAGTTCCTGCTTAAACCCCTGATGACGATTTTGACAGCAGGTCTCAA 672
QY 1128 GATTTAAAGCTATTTGAAACCACTAGTATCGGTGACACATCACTCTCAAGATTAATA 1187
DB 673 GATTTAAAGCTATTTGAAACCACTAGTATCGGTGACACATCACTCTCAAGATTAATA 732
QY 1188 GCTCAAGCAAGAAAGATTTTAAACAAACCCAGCGCTATACGATTTATGAAACGTGAC 1247
DB 733 GCTCAAGCAAGAAAGATTTTAAACAAACCCAGCGCTATACGATTTATGAAACGTGAC 792
QY 1248 TCCCTCAATCGTCACTCATGACATGACATTTTCCGATGCAATTTTACCAATGATCAAGAG 1307
DB 793 TCCCTCAATCGTCACTCATGACATGACATTTTCCGATGCAATTTTACCAATGATCAAGAG 852
QY 1308 TTTACTTACCGTGTAAATAATCGGGAACCAAGCTTATAGATCAATAAATAATCGTGTCTG 1367
DB 853 TTTACTTACCGTGTAAATAATCGGGAACCAAGCTTATAGATCAATAAATAATCGTGTCTG 912
QY 1368 AATGAAGAAATAAACAACGACCTGATCTTGAAATAATTAAGTCTCTTAAATAAAGGG 1427
DB 913 AATGAAGAAATAAACAACGACCTGATCTTGAAATAATTAAGTCTCTTAAATAAAGGG 972
QY 1428 GAAAAGCCGTATGATCCCTTTGATGCGAGTCACTTGAATCTGTTCAACATCAATAGCTT 1487
DB 973 GAAAAGCCGTATGATCCCTTTGATGCGAGTCACTTGAATCTGTTCAACATCAATAGCTT 1032
QY 1488 GATGTCGATACCAAGCAATTTGCTAAATAAGTGAAGACGCTTTAACAGCTACGGAACGTAAC 1547
DB 1033 GATGTCGATACCAAGCAATTTGCTAAATAAGTGAAGACGCTTTAACAGCTACGGAACGTAAC 1092
QY 1548 TTAGACTTCAAGATTTTATACATCTCGTGAATAGGCTAACTACTTACCAACAAATCTC 1607

```

```

DB 1093 TTAGACTTCAAGATTTTATACATCTCGTGAATAGGCTAACTACTTACCAACATCTC 1152
QY 1608 GATGCTTTTGGTATTTATGACATATACCTTAACTGGAATAAGTGAATATCAGATGAC 1667
DB 1153 GATGCTTTTGGTATTTATGACATATACCTTAACTGGAATAAGTGAATATCAGATGAC 1212
QY 1668 ACCAAGCTATTCATACCGTTTATATGAGGACGACCCGAAGGAGGAATGCTAGACTAC 1727
DB 1213 ACCAAGCTATTCATACCGTTTATATGAGGACGACCCGAAGGAGGAATGCTAGACTAT 1272
QY 1728 CATTAGCTGATGAT 1742
DB 1273 CATTAGCTATGAT 1287

```

RESULT 6 US-10-631-558-1

```

; Sequence 1, Application US/10631558
; Publication No. US20050260598A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kamara
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/10/631,558
; PRIOR FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US/09/940,235
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1245
; TYPE: DNA
; ORGANISM: Streptococcus equisimilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1242)
US-10-631-558-1

```

Query Match 54.6%; Score 1145.4; DB 7; Length 1245;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1149; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

QY 588 ATTGCTGAGCCTGAGGCGTCTAGACCGTCCATCTGTGACAAACAGCCAAATGGTGT 647
DB 1 ATTGCTGAGCCTGAGGCGTCTAGACCGTCCATCTGTGACAAACAGCCAAATGGTGT 60
QY 648 AGCGTTGCTGCTACTGTTGAGGGGAGCAATCAAGACTTATGTTAAATTTTGAATC 707
DB 61 AGCGTTGCTGCTACTGTTGAGGGGAGCAATCAAGACTTATGTTAAATTTTGAATC 120
QY 708 GATCTAACATCAAGACCTGCTCATGAGGAGAAAGACAGCAAGCTTAAAGTCCAAATCA 767
DB 121 GATCTAACATCAAGACCTGCTCATGAGGAGAAAGACAGCAAGCTTAAAGTCCAAATCA 180
QY 768 AAACCAATTTGCTACTGATAGTGGCGGATGTCAATAACTTGAGAAAGTGACTTACTA 827
DB 181 AAACCAATTTGCTACTGATAGTGGCGGATGTCAATAACTTGAGAAAGTGACTTACTA 240
QY 828 AAGGCTATTCAAGAACAAATTTGATCGCTAACGTCCACAGTAAAGCACTACTTGAAGTC 887

```



```

Db      ||||| 241 AAGGCTATTTCAGAACAAATGATGCTMAAGCTCCACGTAACGACGCTACTTTGAGAGTC 300
Qy      ||||| 888 ATTGATTTTGGCAAGCAGATGCAACATTAATGATGAAAGGCAAGCTACTTTGCTGAC 947
Db      ||||| 301 ATTGATTTTGGCAAGCAGATGCAACATTAATGATGAAAGGCAAGCTACTTTGCTGAC 360
Qy      ||||| 948 AAGATGTTGCTGTAACCTTGCCGACCCAACTGCTCCAGAAATTTTGTCTAAGCGGACAT 1007
Db      ||||| 361 AAGATGTTGCTGTAACCTTGCCGACCCAACTGCTCCAGAAATTTTGTCTAAGCGGACAT 420
Qy      ||||| 1008 GTGGCGGTTTGAACCATTAATAAGAAAAACAATCAAAAACCAAGGAAATCTGTTGATGTG 1067
Db      ||||| 421 GTGGCGGTTTGAACCATTAATAAGAAAAACAATCAAAAACCAAGGAAATCTGTTGATGTG 480
Qy      ||||| 1068 GAATATACGTGTACAGTTTACTCCCTTAAACCCGTGATGACATTTCAACAGGCTCAAA 1127
Db      ||||| 481 GAATATACGTGTACAGTTTACTCCCTTAAACCCGTGATGACATTTCAACAGGCTCAAA 540
Qy      ||||| 1128 GATATACGTGTATTTGAAAAACAATGATGCTGATGACATCAATCTCAAGAAATTAATA 1187
Db      ||||| 541 GATATACGTGTATTTGAAAAACAATGATGCTGATGACATCAATCTCAAGAAATTAATA 600
Qy      ||||| 1188 GCTCAAGCACAAAAGCATTTTAAACAAAACCAACCCAGGCTATACGATTTATGAACGTGAC 1247
Db      ||||| 601 GCTCAAGCACAAAAGCATTTTAAACAAAACCAACCCAGGCTATACGATTTATGAACGTGAC 660
Qy      ||||| 1248 TCTCTAATGCTCATCTGATGACATGACATTTTCCGTAACGATTTTCAATGATGACAG 1307
Db      ||||| 661 TCTCTAATGCTCATCTGATGACATGACATTTTCCGTAACGATTTTCAATGATGACAG 720
Qy      ||||| 1308 TTTACTTACCGTGTAAAAATCGGGAACAAGCTTATAGATCAATAAAAAATCGGTG 1367
Db      ||||| 721 TTTACTTACCGTGTAAAAATCGGGAACAAGCTTATAGATCAATAAAAAATCGGTG 780
Qy      ||||| 1368 AATGAAGAATAAACAACAACCTGATCTCTGAGAAATATTTACGTCCTTAAAAAAGG 1427
Db      ||||| 781 AATGAAGAATAAACAACAACCTGATCTCTGAGAAATATTTACGTCCTTAAAAAAGG 840
Qy      ||||| 1428 GAAAAGCCGTATGATCCCTTTGATCGAGTCACTTGAACCTGTTCAACATCAATACGTT 1487
Db      ||||| 841 GAAAAGCCGTATGATCCCTTTGATCGAGTCACTTGAACCTGTTCAACATCAATACGTT 900
Qy      ||||| 1488 GATGTGATACCAAGCAATTTGCTAAAAAGTGAAGACAGTCTTAAACAGTACGAAACGTAAC 1547
Db      ||||| 901 GATGTGATACCAAGCAATTTGCTAAAAAGTGAAGACAGTCTTAAACAGTACGAAACGTAAC 960
Qy      ||||| 1548 TTAAGCTTCAAGATTTATACGATCCCTGATGATAGGCTTAACTACTCTCAACAATCTC 1607
Db      ||||| 961 TTAAGCTTCAAGATTTATACGATCCCTGATGATAGGCTTAACTACTCTCAACAATCTC 1020
Qy      ||||| 1608 GATGCTTTTGTATTAATGAATTAATACCTTAACCTGAAAAAGTAGAGATTAATCAAGATAC 1667
Db      ||||| 1021 GATGCTTTTGTATTAATGAATTAATACCTTAACCTGAAAAAGTAGAGATTAATCAAGATAC 1080
Qy      ||||| 1668 ACCAACCGTATCAATAACGTTTATATAGGGAAGCGAACCCGAAGAGAAATGCTAGTAC 1727
Db      ||||| 1081 ACCAACCGTATCAATAACGTTTATATAGGGAAGCGAACCCGAAGAGAAATGCTAGTAC 1140
Qy      ||||| 1728 CATTTAGCTGTGTGT 1742
Db      ||||| 1141 CATTTAGCTGTGTAT 1155

```

RESULT 7

US-10-631-558-6

; Sequence 6, Application US/10631558

; Publication No. US20050260598A1

; GENERAL INFORMATION:

; APPLICANT: Kumar, Rajesh

; APPLICANT: Sahni, Girish

; APPLICANT: Roy, Chait

; APPLICANT: Rajagopal, Kammarra

```

; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OR INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/10/631,558
; PRIOR APPLICATION NUMBER: US/09/940,235
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1327
; TYPE: DNA
; ORGANISM: Streptococcus equisimilis
; US-10-631-558-6

Query Match      54.1%; Score 1134.2; DB 7; Length 1327;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1142; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy      ||||| 588 ATTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCAATGGTTGT 647
Db      ||||| 83 ATAGCTGCTCTGAATGGCTATAGATGCTCTTCTGTAATAACAGCAATGGTTGT 142
Qy      ||||| 648 AGCTGTCTGTACTGTGTTAGAGGGAAGCAATCAAGACATTAATTTTGAATC 707
Db      ||||| 143 AGCTGTCTGTACTGTGTTAGAGGGAAGCAATCAAGACATTAATTTTGAATC 202
Qy      ||||| 708 GATTAACATCAAGACCTGCTCATGAGGAAGAAGCAAGCAAGGCTTAAGTCCAAATCA 767
Db      ||||| 203 GATTAACATCAAGACCTGCTCATGAGGAAGAAGCAAGCAAGGCTTAAGTCCAAATCA 262
Qy      ||||| 768 AAACATTTGCTACTGTAAGTGGCGCATGTCATATAACTTTGAAAAAGTGACTTAATA 827
Db      ||||| 263 AAACATTTGCTACTGTAAGTGGCGCATGTCATATAACTTTGAAAAAGTGACTTAATA 322
Qy      ||||| 828 AAGCTATTCAGAAACATTAATGATGCTTAAGTCAAGTCAAGTCACTTTGAGGTC 887
Db      ||||| 323 AAGCTATTCAGAAACATTAATGATGCTTAAGTCAAGTCAAGTCACTTTGAGGTC 382
Qy      ||||| 888 ATTGATTTTGAAGCGATGCAACATTAATGATGGAAGGCAAGGCTACTTTGCTGAC 947
Db      ||||| 383 ATTGATTTTGAAGCGATGCAACATTAATGATGGAAGGCAAGGCTACTTTGCTGAC 442
Qy      ||||| 948 AAGATGTTTGGTAACTTGGCCGACCCAACTGTCCAAAGAAATTTTGTAAAGGCAAT 1007
Db      ||||| 443 AAGATGTTTGGTAACTTGGCCGACCCAACTGTCCAAAGAAATTTTGTAAAGGCAAT 502
Qy      ||||| 1008 GTGGCGGTTTGAACCATTAATAAGAAAAACAATCAAAAACCAAGGAAATCTGTTGATGTG 1067
Db      ||||| 503 GTGGCGGTTTGAACCATTAATAAGAAAAACAATCAAAAACCAAGGAAATCTGTTGATGTG 562
Qy      ||||| 1068 GAATATACGTGTACAGTTTACTCCCTTAAACCCGTGATGACATTTCAACAGGCTCAAA 1127
Db      ||||| 563 GAATATACGTGTACAGTTTACTCCCTTAAACCCGTGATGACATTTCAACAGGCTCAAA 622
Qy      ||||| 1128 GATATACGTGTATTTGAAAAACAATGATGCTGATGACATCAATCTCAAGAAATTAATA 1187
Db      ||||| 623 GATATACGTGTATTTGAAAAACAATGATGCTGATGACATCAATCTCAAGAAATTAATA 682
Qy      ||||| 1188 GCTCAAGCACAAAAGCATTTTAAACAAAACCAACCCAGGCTATACGATTTATGAACGTGAC 1247
Db      ||||| 683 GCTCAAGCACAAAAGCATTTTAAACAAAACCAACCCAGGCTATACGATTTATGAACGTGAC 742
Qy      ||||| 1248 TCTCTAATGCTCATCTGATGACATGACATTTTCCGTAACGATTTTCAATGATGACAG 1307

```

```
Db 743 TCCATATGTCATCATGACATGACATTTTCGATACGTTTACCAATGATCAAG 802
Qy 1308 TTTACTACCGTGTAAATATGGGAAACAAGCTTATAGATCAATATATGATGCTG 1367
Db 803 TTTACTACCGTGTAAATATGGGAAACAAGCTTATAGATCAATATATGATGCTG 862
Qy 1368 AATGAAGAATTAACAACATGCTGATCTGTAGAAATATATAGCTTAAATGAGG 1427
Db 863 AATGAAGAATTAACAACATGCTGATCTGTAGAAATATATAGCTTAAATGAGG 922
Qy 1428 GAAAACCGTATGATCCCTTTGATGCGAGTCACTTGAATCTTCAACATCAATGCT 1487
Db 923 GAAAACCGTATGATCCCTTTGATGCGAGTCACTTGAATCTTCAACATCAATGCT 982
Qy 1488 GATGTCGATACCAACCAATTTGCTTAAATATGAGACACTCTTAAACAGTACGAACTG 1547
Db 983 GATGTCGATACCAACCAATTTGCTTAAATATGAGACACTCTTAAACAGTACGAACTG 1042
Qy 1548 TTAGACTTGAAGATTTATACGATCCTCGTATAGGCTTAACTACTCTTACCAATCTG 1607
Db 1043 TTAGACTTGAAGATTTATACGATCCTCGTATAGGCTTAACTACTCTTACCAATCTG 1102
Qy 1608 GATGCTTTGTATTATATGACTATACCTTAACTGAAAATGAGATATCAAGATGAC 1667
Db 1103 GATGCTTTGTATTATATGACTATACCTTAACTGAAAATGAGATATCAAGATGAC 1162
Qy 1668 ACCAACCGTATCAATACCGTTTATATGCGCAAGCAACCGAGAGAGAAATGCTAGTAC 1727
Db 1163 ACCAACCGTATCAATACCGTTTATATGCGCAAGCAACCGAGAGAGAAATGCTAGTAC 1222
Qy 1728 CATTTAGCTGNGGT 1742
Db 1223 CATTTAGCTGNGGT 1237
```

RESULT 8

```
US-10-631-558-3
; Sequence 3, Application US/10631558
; Publication No. US20050260598A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammarra
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLON-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/10/631,558
; CURRENT FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US/09/940,235
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 777
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1) ... (777)
US-10-631-558-3
```

Query Match 15.7%; Score 330; DB 7; Length 777;

Best Local Similarity 100.0%; Pred. No. 4.3e-90;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1764 CCATAGCTGAGAGAGTGTATTGATCATGCTGCTGGAGCTTCCATATGTGTGCGAGAAACG 1823
Db 448 CCATAGCTGAGAGAGTGTATTGATCATGCTGCTGGAGCTTCCATATGTGTGCGAGAAACG 507
Qy 1824 TGGAGAGAGCCCTACCAAGGCTGATGATGTGTATGTAATTGACTTGGAGAGAGGAGC 1883
Db 508 TGGAGAGAGCCCTACCAAGGCTGATGATGTGTATGTAATTGACTTGGAGAGAGGAGC 567
Qy 1884 GAGCATCACTTGACCTTTAGAAATATAGATGCAACATGAGACACAAGACATCTAT 1943
Db 568 GAGCATCACTTGACCTTTAGAAATATAGATGCAACATGAGACACAAGACATCTAT 627
Qy 1944 AGAATTGAGACACCTGAGCAAGAGATATGAGAGAAACCTGCTCCAGTATCTGC 2003
Db 628 AGAATTGAGACACCTGAGCAAGAGATATGAGAGAAACCTGCTCCAGTATCTGC 687
Qy 2004 ACAGGCAACGCGGAGAGAGTGAAGTGTGAGAGGACACCTGTGAGAGACCAATCG 2063
Db 688 ACAGGCAACGCGGAGAGAGTGAAGTGTGAGAGGACACCTGTGAGAGACCAATCG 747
Qy 2064 AGCGATCTGCGCCCTTCAACGATGTTCT 2093
Db 748 AGCGATCTGCGCCCTTCAACGATGTTCT 777
```

RESULT 9

```
US-10-995-561-114
; Sequence 114, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 114
; LENGTH: 2443
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-114
```

Query Match 15.7%; Score 330; DB 7; Length 2443;
Best Local Similarity 100.0%; Pred. No. 8.4e-90;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1764 CCATAGCTGAGAGAGTGTATTGATCATGCTGCTGGAGCTTCCATATGTGTGCGAGAAACG 1823
Db 913 CCATAGCTGAGAGAGTGTATTGATCATGCTGCTGGAGCTTCCATATGTGTGCGAGAAACG 972
Qy 1824 TGGAGAGAGCCCTACCAAGGCTGATGATGTGTATGTAATTGACTTGGAGAGAGGAGC 1883
Db 973 TGGAGAGAGCCCTACCAAGGCTGATGATGTGTATGTAATTGACTTGGAGAGAGGAGC 1032
Qy 1884 GAGCATCACTTGACCTTTAGAAATATAGATGCAACATGAGACACAAGACATCTAT 1943
Db 1033 GAGCATCACTTGACCTTTAGAAATATAGATGCAACATGAGACACAAGACATCTAT 1092
Qy 1944 AGAATTGAGACACCTGAGCAAGAGATATGAGAGAAACCTGCTCCAGTATCTGC 2003
Db 1093 AGAATTGAGACACCTGAGCAAGAGATATGAGAGAAACCTGCTCCAGTATCTGC 1152
Qy 2004 ACAGGCAACGCGGAGAGAGTGAAGTGTGAGAGGACACCTTGTGAGACCAATCG 2063
Db 1153 ACAGGCAACGCGGAGAGAGTGAAGTGTGAGAGGACACCTTGTGAGACCAATCG 1212
Qy 2064 AGCGATCTGCGCCCTTCAACGATGTTCT 2093
```


Qy 2064 AGCGATCTGGCCCCCTTCACCGATGTTGCT 2093
|||
Db 1213 AGCGATCTGGCCCCCTTCACCGATGTTGCT 1242

Search completed: February 1, 2006, 14:43:15
Job time : 387.523 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2006, 16:09:44 ; Search time 8612.84 Seconds
(without alignments)
11385.988 Million cell updates/sec

Title: US-09-940-235-12

Perfect score: 2096
Sequence: 1 cgaagaccatcatcgtgtt.....cctcacgcatgtcgttag 2096

Scoring table: IDENTITY_NTC
Gapop 10.0, Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	330	15.7	451	1	AL603368 DKFZp686C
2	330	15.7	560	7	CN419585 170005318
3	330	15.7	621	7	CN482442 h200008.Y
4	330	15.7	626	7	CN419611 170005313
5	330	15.7	737	1	AU140993 AU140993
6	330	15.7	765	3	BM715855 UI-B-EJ0-
7	330	15.7	861	1	AU140971 AU140971
8	330	15.7	7501	4	BC078656 Homo sapi
9	330	15.7	7501	4	BC100030 Homo sapi
10	330	15.7	7777	4	CR749316 Homo sapi
11	330	15.7	7868	4	HSW80653 Homo sapi
12	330	15.7	7885	4	CR749281 Homo sapi
13	330	15.7	8411	4	HSW806992 Homo sapi
14	329	15.7	465	1	AI095589 CB24a08.X
15	329	15.7	734	1	AU140802 AU140802
16	328.4	15.7	8121	4	CR749317 Homo sapi
17	326.8	15.6	548	6	CD613789 56022208J
18	325.8	15.6	551	6	CD613788 56022208J
19	325.2	15.5	560	1	AI743013 WGSa12.X
20	319	15.2	500	2	BP956982 OVA-NN114
21	318.2	15.2	547	2	BE009640 PM4-BN017
22	318	15.2	943	5	BX391752 BX391752

23	317	15.1	705	1	AU140789
24	315.4	15.0	739	1	AL706215
25	310	14.8	603	7	CN419658
26	307	14.6	414	1	AA492032
27	305	14.6	518	1	AI093548
28	304	14.5	7434	11	DQ039102
29	298	14.2	538	1	AL603599
30	297.4	14.2	496	2	BG945197
31	296.4	14.1	303	1	AA852090
32	287	13.7	427	7	CN419650
33	287	13.7	452	7	CN419649
34	281	13.4	605	2	CN419657
35	280	13.4	408	2	BG900107
36	277.8	13.3	580	3	BQ292415
37	274	13.1	495	2	BP956977
38	271.8	13.0	494	2	BP956365
39	268.6	12.8	484	2	B1058354
40	268	12.8	725	5	BQ574857
41	264	12.6	454	8	W46530
42	261.8	12.5	474	6	CB536952
43	256.8	12.3	645	8	DN123922
44	251.8	12.0	314	7	CN419631
45	249.4	11.9	899	5	BX327266

ALIGNMENTS

RESULT 1
AL603368 451 bp mRNA linear EST 04-SEP-2003
LOCUS DKFZp686C067.5, mRNA sequence.
DEFINITION DKFZp686C067.5, mRNA sequence.
ACCESSION AL603368
VERSION AL603368.1 GI:15166874
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
TITLE This is the 5' sequence of the clone insert
JOURNAL Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
COMMENT Reseach Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by EMBL (European Molecular Biology Laboratories,
Heidelberg/Germany) within the CDNA sequencing consortium of the
German Genome Project.
NO 81 sequence available.
This clone (DKFZp686C067) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcententrum, Hubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Contact: MIPS
MIPS
location/Qualifiers

FEATURES

source
1..451
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686C067"
/dev_stage="adult"
/lab_host="DH10B"
/clone_id="686 (synonym: hlc3)"
/note="Vector: pTriblEx2; Site_1: SfiIA; Site_2: SfiIB;
CDNA-collection"

ORIGIN

Query Match 15.7%; Score 330; DB 1; Length 451;
Best Local Similarity 100.0%; Pred. No. 8.4e-84;

Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1764 CCCATAGCTGAGAAGTGTGTTGATCATGCTGCTGGACCTTCTATGTGTGCGAAGAACG 1823
 |||||
 Db 32 CCCATAGCTGAGAAGTGTGTTGATCATGCTGCTGGACCTTCTATGTGTGCGAAGAACG 91
 |||||

QY 1824 TGGAGAAAGCCCTTACCAAGGCTGATGATGTAATTGTACTTCCCTGGAGAGGACG 1883
 |||||
 Db 92 TGGAGAAAGCCCTTACCAAGGCTGATGATGTAATTGTACTTCCCTGGAGAGGACG 151
 |||||

QY 1884 GGAAGCATCACTTGCACCTTTAGAAATGATGACATGACACACAGACATCTTAT 1943
 |||||
 Db 152 GGAAGCATCACTTGCACCTTTAGAAATGATGACATGACACACAGACATCTTAT 211
 |||||

QY 1944 AGAATTGAGACACCTGAGACAGAAAGATATGAGAAACCTGCTCCAGTGCATCTGC 2003
 |||||
 Db 212 AGAATTGAGACACCTGAGACAGAAAGATATGAGAAACCTGCTCCAGTGCATCTGC 271
 |||||

QY 2004 ACAGGCAACGCGCCGAGAGAGTGAAGTGTAGAGGACACCTCTGTGACAGACATCG 2063
 |||||
 Db 272 ACAGGCAACGCGCCGAGAGAGTGAAGTGTAGAGGACACCTCTGTGACAGACATCG 331
 |||||

QY 2064 AGCGGATCTGGCCCTTCAACCGATGTTCTG 2093
 |||||
 Db 332 AGCGGATCTGGCCCTTCAACCGATGTTCTG 361
 |||||

RESULT 2
 LOCUS CN419585 560 bp mRNA linear EST 16-MAY-2004
 DEFINITION 17000531856084 GRN_EB Homo sapiens CDNA 5', mRNA sequence.
 ACCESSION CN419585
 VERSION CN419585.1 GI:47407179
 KEYWORDS EST.

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 560)
 Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,
 Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R.,
 Lebowitz, J. and Stanton, L.W.
 Transcriptional characterization elucidates signaling networks that
 control human ES cell growth and differentiation
 Nat. Biotechnol. 22 (6), 707-716 (2004)

TITLE JOURNAL
 PUBMED
 COMMENT

Contact: Brandenberger R
 Regenerative Medicine
 Geron Corporation
 230 Constitution Drive, Menlo Park, CA 94025, USA
 Tel: 650 473 8658
 Fax: 650 473 7760
 Email: rbrandenberger@geron.com
 Insert Length: 560 Std Error: 0.00.

FEATURES
 source location/Qualifiers
 1..560

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /tissue_type="embryonic stem cells, embryoid bodies
 derived from H1, H7 and H9 cells"
 /clone_lib="GRN_EB"
 /note="oligo dt primed, full-length enriched cDNA library
 from embryoid body outgrowths derived from hES cell lines
 H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free
 conditions."

ORIGIN

Query Match 15.7%; Score 330; DB 7; Length 560;
 Best Local Similarity 100.0%; Pred. No. 9e-84;
 Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1764 CCCATAGCTGAGAAGTGTGTTGATCATGCTGCTGGACCTTCTATGTGTGCGAAGAACG 1823
 |||||
 Db 209 CCCATAGCTGAGAAGTGTGTTGATCATGCTGCTGGACCTTCTATGTGTGCGAAGAACG 268
 |||||

QY 1824 TGGAGAAAGCCCTTACCAAGGCTGATGATGTAATTGTACTTCCCTGGAGAGGACG 1883
 |||||
 Db 269 TGGAGAAAGCCCTTACCAAGGCTGATGATGTAATTGTACTTCCCTGGAGAGGACG 328
 |||||

QY 1884 GGAAGCATCACTTGCACCTTTAGAAATGATGACATGACACACAGACATCTTAT 1943
 |||||
 Db 329 GGAAGCATCACTTGCACCTTTAGAAATGATGACATGACACACAGACATCTTAT 388
 |||||

QY 1944 AGAATTGAGACACCTGAGACAGAAAGATATGAGAAACCTGCTCCAGTGCATCTGC 2003
 |||||
 Db 389 AGAATTGAGACACCTGAGACAGAAAGATATGAGAAACCTGCTCCAGTGCATCTGC 448
 |||||

QY 2004 ACAGGCAACGCGCCGAGAGAGTGAAGTGTAGAGGACACCTCTGTGACAGACATCG 2063
 |||||
 Db 449 ACAGGCAACGCGCCGAGAGAGTGAAGTGTAGAGGACACCTCTGTGACAGACATCG 508
 |||||

QY 2064 AGCGGATCTGGCCCTTCAACCGATGTTCTG 2093
 |||||
 Db 509 AGCGGATCTGGCCCTTCAACCGATGTTCTG 538
 |||||

RESULT 3
 LOCUS CN482442 621 bp mRNA linear EST 26-APR-2004
 DEFINITION hw20d08.y1 Human primary human ocular pericytes. Unamplified (hw)
 Homo sapiens CDNA clone hw20d08 5', mRNA sequence.
 ACCESSION CN482442
 VERSION CN482442.1 GI:46563946
 KEYWORDS EST.

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 621)
 Teal, J.Y. and Wistow, G.
 Expressed sequence tag analysis of cultured primary human ocular
 pericytes
 Unpublished (2004)

JOURNAL
 PUBMED
 COMMENT

Contact: Wistow G
 Section on Molecular Structure and Function
 National Eye Institute
 6/331, NIH, Bethesda, MD 20892-2740, USA
 Tel: 301 402 3452
 Fax: 301 496 0078
 Email: graeme@helix.nih.gov
 Plate: 20 row: d column: 08
 Seq primer: M13RP1 reverse primer (ABI).

FEATURES
 source location/Qualifiers
 1..621

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="hw20d08"
 /cell_type="pericytes"
 /dev_stage="Adult"
 /lab_host="EMDH108"
 /clone_lib="Human primary human ocular pericytes.
 Unamplified (hw)"

/note="Organ: Eye; Vector: pSport1; RNA was extracted from
 primary human pericytes in culture. A directionally cloned
 cDNA library in the pSPORT1 vector (Invitrogen) was
 constructed at Bioserve Biotechnology (Laurel MD)
 essentially following the protocols of the SuperScript
 Plasmid System full details of which are contained in the
 manufacturer's instruction manual
 (http://www.lifetechn.com/). First strand synthesis was
 carried out using a Not I primer-adaptor
 [5'-pACTAGTCTTAGATCGAGCGGCCCC(TT)5-3']. cDNA was

Query Match 15.7%; Score 330; DB 1; Length 737;
 Best Local Similarity 100.0%; Pred. No. 9.8e-84;
 Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1764 CCCATAGCTGAGAGAGTTTGTATCATGCTGCTGGACCTTCTATGTGTGGAGAAACG 1823
 |||||
 DB 362 CCCATAGCTGAGAGAGTTTGTATCATGCTGCTGGACCTTCTATGTGTGGAGAAACG 421

QY 1824 TGGGAGAGAGCCCTACCAAGGCTGATGATGATGATTTGATCTTGGCCCTGGAGAAAGGAGC 1883
 |||||
 DB 422 TGGGAGAGAGCCCTACCAAGGCTGATGATGATGATTTGATCTTGGCCCTGGAGAAAGGAGC 481

QY 1884 GAGCGCATCACTTGCATCTTCTAGAAATAGATGACATGACAGACACAAAGACATCTTAT 1943
 |||||
 DB 482 GAGCGCATCACTTGCATCTTCTAGAAATAGATGACATGACAGACACAAAGACATCTTAT 541

QY 1944 AGAATTGGAGACACCTGGAGCAAGAAAGATATGACAGAAACCTGCTCCAGTGATCTGC 2003
 |||||
 DB 542 AGAATTGGAGACACCTGGAGCAAGAAAGATATGACAGAAACCTGCTCCAGTGATCTGC 601

QY 2004 ACAGGCAAGGCGCGAGAGAGTGGAGTGGAGAGGACACCTCTGTGACAGACACATCG 2063
 |||||
 DB 602 ACAGGCAAGGCGCGAGAGAGTGGAGTGGAGAGGACACCTCTGTGACAGACACATCG 661

QY 2064 AGCGGATCTGGCCCTTCAACCGATTTCT 2093
 |||||
 DB 662 AGCGGATCTGGCCCTTCAACCGATTTCT 691

RESULT 6 765 bp mRNA linear EST 28-FEB-2002
 LOCUS BM715855
 DEFINITION UI-E-BJ0-abj-h-11-0-UI.r2 UI-E-BJ0 Homo sapiens cDNA clone
 LOCUS BM715855
 VERSION BM715855.1 GI:19029113
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 1 (bases 1 to 765)
 Ronaldo, M.F., Lennon, G. and Soares, M.B.
 Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 PUBMED 8889548
 COMMENT Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel. 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@iowa.edu
 Tissue Procurement: Dr. Gregg Hageman
 cDNA library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.reagen.com).
 Seq primer: M13 Reverse.
 FEATURES
 source
 1..765
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-E-BJ0-abj-h-11-0-UI"
 /tissue_type="fetal eyes, lens, eye anterior segment,
 optic nerve, retina, Retina Foveal and Macular, RPE and
 Choroid"
 /dev_stage="fetal and adult"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"

Query Match 15.7%; Score 330; DB 3; Length 765;
 Best Local Similarity 100.0%; Pred. No. 9.9e-84;
 Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1764 CCCATAGCTGAGAGAGTTTGTATCATGCTGCTGGACCTTCTATGTGTGGAGAAACG 1823
 |||||
 DB 182 CCCATAGCTGAGAGAGTTTGTATCATGCTGCTGGACCTTCTATGTGTGGAGAAACG 241

QY 1824 TGGGAGAGAGCCCTACCAAGGCTGATGATGATGATTTGATCTTGGCCCTGGAGAAAGGAGC 1883
 |||||
 DB 242 TGGGAGAGAGCCCTACCAAGGCTGATGATGATGATTTGATCTTGGCCCTGGAGAAAGGAGC 301

QY 1884 GAGCGCATCACTTGCATCTTCTAGAAATAGATGACATGACAGACACAAAGACATCTTAT 1943
 |||||
 DB 302 GAGCGCATCACTTGCATCTTCTAGAAATAGATGACATGACAGACACAAAGACATCTTAT 361

QY 1944 AGAATTGGAGACACCTGGAGCAAGAAAGATATGACAGAAACCTGCTCCAGTGATCTGC 2003
 |||||
 DB 362 AGAATTGGAGACACCTGGAGCAAGAAAGATATGACAGAAACCTGCTCCAGTGATCTGC 421

QY 2004 ACAGGCAAGGCGCGAGAGAGTGGAGTGGAGAGGACACCTCTGTGACAGACACATCG 2063
 |||||
 DB 422 ACAGGCAAGGCGCGAGAGAGTGGAGTGGAGAGGACACCTCTGTGACAGACACATCG 481

QY 2064 AGCGGATCTGGCCCTTCAACCGATTTCT 2093
 |||||
 DB 482 AGCGGATCTGGCCCTTCAACCGATTTCT 511

RESULT 7 861 bp mRNA linear EST 05-AUG-2002
 LOCUS AU140971
 DEFINITION AU140971 PLACE4 Homo sapiens cDNA clone PLACE4000583 5', mRNA
 sequence.
 ACCESSION AU140971.1 GI:11002492
 VERSION AU140971.1
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 1 (bases 1 to 861)
 Ota, T., Suzuki, Y., Saito, K., Ishii, S., Yamamoto, J., Sugiyama, T.,
 Nishikawa, T., Nakamura, Y., Sugano, S., Masuno, Y. and Isogai, T.,
 HRI human cDNA project (Ota, T., Suzuki, Y., Saito, K., Ishii, S.,
 Yamamoto, J., Sugiyama, T., Nishikawa, T., Nakamura, Y., Sugano, S.,
 Masuno, Y., Isogai, T.)
 Unpublished (2000)
 Contact: Takao Isogai
 Genomics Laboratory
 Helix Research Institute
 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan

/clone lib="UI-E-BJ0"
 /note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
 modified polylinker; Site 1: EcoR I; Site 2: Not I;
 UI-E-BJ0 is a subtracted cDNA library constructed
 according to Ronaldo, Lennon and Soares, Genome Research,
 6:791-806, 1996. First strand cDNA synthesis was primed
 with an oligo-dT primer containing a Not I site. Double
 stranded cDNA was ligated to an EcoR I adaptor, digested
 with Not I, and cloned directionally into pT73-Pac
 vector. The oligonucleotide used to prime the synthesis of
 first-strand cDNA contains a library tag sequence that is
 located between the Not I site and the (dT)18 tail. The
 sequence tags for this library are: fetal eyes,
 AGAATGACAG; lens, CGATTAGGCA; eye anterior segment,
 AATGCCGAT; optic nerve, CCATTAGAG; retina, CCGCG;
 Foveal and Macular, GTCC; RPE and Choroid, ACTTA. This
 library was created for the program, Gene Discovery in the
 Visual System, supported by National Eye Institute (NEI)."


```
QY      2064 AGCGAATCTGGCCCCCTTCAACGATGTTCGT 2093
        |||||
Db      1106 AGCGAATCTGGCCCCCTTCAACGATGTTCGT 1135
```

RESULT 9				
BC100030				
LOCUS	BC100030	7501 bp	mRNA	linear
DEFINITION	Homo sapiens	cdna clone IMAGE:30343662,	containing	frame-shift

/clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host
DH10B; sites SfiIA + SfiIB"
/dev_stage="adult"
/note="fibronectin 1, differentially spliced"
1..7777
/gene="DKFZp686K08164"
266..7069
/gene="DKFZp686K08164"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAH18171.1"
/db_xref="GI:51476362."
/translation="MLRGPGGLLLAVLCITGAVPSTGASKSKRAQNVQPSVA
VQSXKGCYDNGKHVQINQWERTYLGALVCTCYGSGRPNCSKPEAEETFDKXT
GATYRVGDTYERKDSIMIDCTCIGARGISCTIARCHEGQSYKIGPTMRPHET
GGYMLEVCLGKNGKEMTCKPIAEKCPDHAAGSYVVGEMWPDYOGMMWVDCICGE
GSGRTCTSRNRCNDODTRTSYRIGPTWSKDRGNLQCTICNGNGEKCRHITSY
QTTSSGSGPPTDVAAYQPPHQPPTGCVTDSGVTSVGMQLTKGNQMLCT
CLANGVSCQETAVTQYTGNSNGPCVLPPTVNGRTFVSCITTEGRDGHLMCSTTSY
BODOKYSFCTDHTVLVOTRGNSNGALCHFPPLNNHNTDCTSEGRDNMKCGTQ
INDADQKFCPEMAAHEICTTEGVWYRIGDOMKQDHGMRCCTGVNGGEMTC
IAYSO.RDQCIYDDITVYVNDTPHKKHBEHMLNCTFCGGRGRWCDPVDQDSEST
GTFYQIDSGMEKHYHVRVQCTCYGKIGKHWQPIQTYPSSSGPPEVPTTPESQPN
SHPIQNAQPSHISKYILKMRPKNSVGKWKATPGLHSYTIKGLKGVVYEGQLI
S1QYGHQVTRPDTFTTSTPVTNTVNTGETTPSPVATSESYTEIATASFVWS
VASADTVSGFRVYELSEBEGDPOYLDPESTASVNIPLLPRKYIVNYQISEDB
OSLISTQTTADAPDPDQVDDTSIVRSRQAPITGRIVYSPESGSETEL
NI.PETANSYTLSDIQVOYONITIVAEENGESTPVIOEETGTPRSQDVPSPRDQ
PFEVDVKTITMTPPESATGYRVAIVPMLGEGKQRIPIKRNFAETVAGISPGYT
YFKVFAVSHGRESKPLTAQOTTKLAPNLQVNSTDVLVAKWMPKQAIQYRLT
VGLTRGQPROVNVGSVSKYPLNLQPSSEVYSVAIGNDSBPATQVFTLLQPG
SSIPNTVETETITVITWTPARIGFKLGRPSOGGAAREVDSGSIIVGSLTQ
GVYVYTIQVLRDQGERDAPLVNKVVTPLSPNTLHLEANDTCVLTYSWERSPTPT
GVYITTPNGQGNSELEVAHDOSSCTFDNISRLGLTVNSYTTKDKESPISDT
IIEVQQLTDLSDVDTDSIGLRTPLNSSTIIGRTIVVAAEGEPIPEDVDSV
GYTYVGLBPGIDYDIDSVITLLNGESAPVLTQOAVPEPTDLRFNIGPDMRWV
APPSIDLTNPLVRSVPVKNBEVAELISPDNAVALNLLGTEVVSVSVECH
ESTPLRGQKTGADSPGIDFSDITANSFVHWIAPRATITGRIRHHPHEDV
DVPSPSRNSITLTLNLTGTEVYVSIVANSNGRESPLIGQOSVSPVPRDLVYAAR
TSLLISWDAVATVRYRTTGTGEGNSPQOETTVGSKSTATISGLKREVDYITVY
ATVGRDSSPASSPISINRTEIDKQSQOVVDVQNSISVKMLPSSSPQVGRVTT
PANGPPTKTAGDPQDQEMTLEGLQPTVEYVSVAQPSGESQPLVQAVTITPAP
TDLKFTQVPTSLSAQMPBNVQLTGYRVAIVPEKTEGPKKEINLAPDSSVVSGLM
VANKVEVSVALKDLTSRPAQGVTTLEVSPRARAVTDATEITITISMRKTEI
TGEQDVAVPANGQPIORTIKDPVRSYTTGTG.OPTDXYKIYTLTADNARSVVIDA
SFAIDAPSLURFATTPTNLSLWSQPPARITTYIKIEKPSPPREVPREPGVTE
AATTGEBETEYTIIVIALKNNQSKBPLGRKKTQGBALSQTTISWAPDQTEYIIS
CHPVGDEBPLQFRVPGTSTATLGTGTGATYNIIVEALKDQHRKHVEVTVGS
VNEGLOPPTDSCFDYTVSHYAVGDEMSEBSGFKLCOCLGFGSGHFRCDSSRMC
HNGVNYKIGEKMDROGNGOMMSCTCLANGKEFKCDPHEATCYDDGKTYHGEQMO
KEYLGAICCTCFGCGRGRCNCRPGRSGBPTEGTGGSYNOYSORHYORTVNVNC
PIECFMPDVOADREDSRE"

ORIGIN

Query Match 15.7%; Score 330; DB 4; Length 7777;
Best Local Similarity 100.0%; Pred. No. 2.1e-83;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1764 CCCATAGCTGAGAGAGTGTTCATGCTGCTGCGGACCTTCATGTGTCGAGAAACG 1823
DB CCCATAGCTGAGAGAGTGTTCATGCTGCTGCGGACCTTCATGTGTCGAGAAACG 865
1824 TGGGAGAGCCCTCAAGGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1883
DB TGGGAGAGCCCTCAAGGCTGATGATGATGATGATGATGATGATGATGATGATGAT 925
1884 GGAACCATCTGCACTTCTGAATPAGATGCAAGATCAGACCAAGGACATCTTAT 1943
DB GGAACCATCTGCACTTCTGAATPAGATGCAAGATCAGACCAAGGACATCTTAT 985
1944 AGAATTGAGACACCTGAGACCAAGAGATTAATCGAGAAAACCTGCTCCAGTGCATCTGC 2003
DB AGAATTGAGACACCTGAGACCAAGAGATTAATCGAGAAAACCTGCTCCAGTGCATCTGC 1045

QY 2004 ACAGGCAACGCGCAGAGAGATGGAAGTGTGAGAGACACCTCTGTGACAGACCATCG 2063
DB 1046 ACAGGCAACGCGCAGAGAGATGGAAGTGTGAGAGACACCTCTGTGACAGACCATCG 1105
QY 2064 AGCGGATCTGGCCCCCTTCACCGATGTTGCT 2093
DB 1106 AGCGGATCTGGCCCCCTTCACCGATGTTGCT 1135

RESULT 11

HSMB06653
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 7868)
Bloecker,H., Boecker,M., Brandt,P., Mewes,H.W., weil,B., Amlid,C.,
Oanger,A., Fobo,G., Han,M. and Wiemann,S.
The German cDNA Consortium
Direct Submission
Submitted (20-JUN-2005) MIPS, Ingolstaedter Landstr.1, D-85764
Neuberberg, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by GBR (National Research Centre for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.
This clone (DKFZp686M04163) is available at the RZPD Deutsches
Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneid=DKFZp686M04163
Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/
Location/Qualifiers

COMMENT

JOURNAL
FEATURES
source
1..7868
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RZPD:DKFZp686M04163Q"
/db_xref="taxon:9606"
/clone="DKFZp686M04163"
/issue_type="uterus endothel, primary cell culture"
/clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host
DH10B; sites SfiIA + SfiIB"
/dev_stage="adult"
/note="fibronectin precursor"
1..7868
/gene="DKFZp686M04163"
267..7157
/gene="DKFZp686M04163"
/codon_start=1
/product="hypothetical protein"
/protein_id="CA845714.1"
/db_xref="GI:34364617"
/db_xref="GCA:Q6N0A6"
/db_xref="InterPro:IPR000083"
/db_xref="InterPro:IPR000562"
/db_xref="InterPro:IPR002086"
/db_xref="InterPro:IPR003961"
/db_xref="InterPro:IPR006209"
/db_xref="UniProt/TREMBL:Q6N0A6"
/translation="MLRGPGGLLLAVLCITGAVPSTGASKSKRAQNVQPSVA
VQSXKGCYDNGKHVQINQWERTYLGALVCTCYGSGRPNCSKPEAEETFDKXT
GATYRVGDTYERKDSIMIDCTCIGARGISCTIARCHEGQSYKIGPTMRPHET
GGYMLEVCLGKNGKEMTCKPIAEKCPDHAAGSYVVGEMWPDYOGMMWVDCICGE
GSGRTCTSRNRCNDODTRTSYRIGPTWSKDRGNLQCTICNGNGEKCRHITSY
QTTSSGSGPPTDVAAYQPPHQPPTGCVTDSGVTSVGMQLTKGNQMLCT

ITGFOVDVAVPANGCPICORTIKFVRSYITTLQPTDVKIYUUTLNDNRSSPVVID
ASTALIDAPSNRLATPNSLWQPPRARITGYI IKKPKSPPRRVRPRRPGVT
BATITGDEPTEYITVIALKNNKSEPLIGKKTDELPLVTLPHNPLPILNDP
STVOKTPEVTHPDIYDNGIOLPDTSGOOSVGOOMI FEEHGRRTPTPTATIRH
PRPYRPNVGOEALISOTTISMAPRODTSSEYIISCHPYGTDEBPLOFRVGRSTATLIG
LIRGATYNIIVALKDQORHRAEYVTVGNSVNEGILNPTDSCDPYIVSHYAGD
EENRSESSEFKLCOCLGFGSGHFRCDSSRWCHNDGNVYKIGKMRQGNQGMWCT
CANGKGRKCDPHEATCYDDGKTVHVGQWQKEYGALCSCTCFGQORMRCDNGCR
PGSRPSEGTGQSYNQSORHQRNTNVNCPICEFMPLDVQADREDSRE"

/mol_type="mRNA"
/db_xref="RZPD:DKFZp68601166q"
/db_xref="taxon:9606"
/clone="DKFZp68601166"
/issue_type="endometrium carcinoma cell line"
/clone_id="686 (synonym: h1c33). Vector pSport1_sfi; host
DH10B; sites SfiIa + SfiIb"
/dev_stage="adult"
/note="fibronectin 1"

Query Match 15.7%; Score 330; DB 4; Length 7885;
Best Local Similarity 100.0%; Pred. No. 2.1e-83;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

gene

1764 CCCATAGCTGAGAGAGTTTGTATCATGCTGCTGGGACCTTCATGTGTCGAGAACG 1823
|||||
846 CCCATAGCTGAGAGAGTTTGTATCATGCTGCTGGGACCTTCATGTGTCGAGAACG 905
|||||
1824 TGGAGAACCCCTACCAAGGCTGATGATGATGATGATGATGATGATGATGATGATG 1883
|||||
906 TGGAGAACCCCTACCAAGGCTGATGATGATGATGATGATGATGATGATGATGATG 965
|||||
1884 GGAAGCATCACTTGCATCTTGAATAATGATGACAGATCAGACACACAGACATCCTAT 1943
|||||
966 GGAAGCATCACTTGCATCTTGAATAATGATGACAGATCAGACACACAGACATCCTAT 1025
|||||
1944 AGAATTTGAGAGACCTTGAGACAAAGATTAATCGAGAAACCTGCTCCAGTGCATTCG 2003
|||||
2004 AGAATTTGAGAGACCTTGAGACAAAGATTAATCGAGAAACCTGCTCCAGTGCATTCG 1085
|||||
2004 ACAGGCAACGCCCGAGAGAGTGAAGTGAAGGACACCTCTGTGAGACACACATCG 2063
|||||
1086 ACAGGCAACGCCCGAGAGAGTGAAGTGAAGGACACCTCTGTGAGACACACATCG 1145
|||||
2064 ACAGGATCTGGCCCTTCAACCGATGTTCTG 2093
|||||
1146 ACAGGATCTGGCCCTTCAACCGATGTTCTG 1175
|||||

CD5
/gene="DKFZp68601166"
268..7701
/gene="DKFZp68601166"
/product="hypothetical protein"
/protein_id="CA845932.1"
/db_xref="GI:34365170"
/db_xref="GOA:Q6MZU5"
/db_xref="InterPro:IPR000083"
/db_xref="InterPro:IPR000562"
/db_xref="InterPro:IPR002086"
/db_xref="InterPro:IPR003961"
/db_xref="InterPro:IPR006209"
/db_xref="UniProt/TREMBL:Q6MZU5"
/translation="MLRGPGLLLAVLCIGTAVPSTGASKSRKROQWVOPSPVA
VSGSKPCYNGHNGHNOQWERTYLGNAIVCTCGSGRPNCSKREARETCDKXT
GMYRVGDTYRERKDSIMDCTCIGARGISCTIARCHEGGSYTGCTWRPHET
GGTMBECVCLANGKENTCKPIABKCTDHAAGSYVGETWERYQGMWVDCCLGE
GSGRTICTSSNRNCDOTRTSYRIGDTWSSKDKRGNLLOCITGNGMRKGCHEMTSV
QTTSSGSGPTDVAAYVQPPHPPQPPYHGSVGVSVGVGMKTOGNKMLCT
CLANGVSCFETAVLYOTGANSNEPCVLPYLYNRRYFSCITRBOGROGCTTSY
EDOKVSCFETAVLYOTGANSNEPCVLPYLYNRRYFSCITRBOGROGCTTSY
NYADOKRGCPMAAHEICTTBGWNRYRGDWDKDHGMRKCTCVNGKERTC
IAYSQLEKQICVDITVNDVTHKREBGMNLCTFEGGKGMKCDPTEQCDQDST
GTFYQIGDSWEKTVHGVRYQCYGRGIGMHCQPLCTFYPSSGPVEFPIEPSON
SHPIOMNAPQPSHISKYLIMRPNKSVGWKEATIPGLHNSYTIKGLKPEVVEGQLI
SIOOYGHOBVTRDFTTSTSTSVTVGERTPSPFLVATSVTEIIVASSVWS
VSAIDTVSGFRVVEIIEEGDEPOYDLPESTASVNPDLIPGRKYVNVYQISEDE
QSILISTQTTAPDAPDPVTDVDTISIVRRSRPOAFTTGRITYSVSBSSTIQ
NLPEIANSVTLSDLOQPOVNYITIVAEENQSTPVTIQETRTGTPSDVPSRDLQ
FVEVTVKTIWMTPEASATGVRVDIPNLGEGHQRPIRMTPEATVTSRGVT
YFKFVAVSHGRESKPLTAQOTKLDLPNLQVNESTVLRMTPEATVTSRGVT
VGLTRGQPOVNVVPSVSKYPLRNPLOPABEYVSVIYIKNGQSPKATGPFVTLQ
SSIPRYNTVETTYITVTPARIGKGVRSQGEARRETVTSBSGSIYVSGLTG
VEYVYITQVLRDQERDAPVNRVVTPLSPNULHEANPDVGLTVSWKSTPDT
GYRIITTTNGQOQNSLEEVQADQSCSTDNISPGLNEVSVYVADDESVISDT
IIEVPQLTDLSEFVDTDSIGRMTPLNSSTIIGYRIIVAAEGEIPFEDVDSV
GYVTVGLBEGIDYDISVTLINGSEAPPTLQOAVPPTDLPFNIGDPNRVTM
APPESIDITNELIVRSPVKNEDVAELISPSMNAVTLTLRGTVEYVSVSVYBOH
ESPLRGRQGTGDSPTGIDSPITANSPTVHMIAPRATITTRIRHHPHSGRPRE
DRVPHSRNSTLTNLTPGTEYVSVIALNREBSPPLIQOSTVSDVDRLEVAAP
TSLISMDAPAVTVRYRITGEGTGSPOBETVPSGKSTATISGLKPEVDTYITV
AYVGRGSPASPKRSINRYTEIDKPSOMTVQONSISKMLPSSPVTGVRVYT
PKNGPQTKTKTAGVTPOTEMTIBLOPVTYVSVYVAAONSSEGSQVVOVATNIBP
KGLAFPDVVDVSIKIMBSPOGOVSRRYVSSPBGIEHLPRAPGEERTALQGR
PGSEIYVSVVVALHDMESOPLTGTSQVSLAPYDLKTOVTVLSAQMPVQVLG
YRVATVPKERTGPMKEINLAPDSSVSATLPAKXEVSVALKOTLTSRPAQVVT
TLENVSPPRARVTDATETITISWRTKSTISCHPYGTDEBPLOFRVGRSTATLIG
YITGLOPDTYKIYUUTLNDNRSSPVVIDASTALIDAPSNRLATPNSLWQPPRARITGYI
PRARITGYIYUUTLNDNRSSPVVIDASTALIDAPSNRLATPNSLWQPPRARITGYI
PLIGRKTDELPLVTLPHNPLPILNDPSTVOKTPEVTHPDIYDNGIOLPDTSGOOSVGOOMI
FEEHGRRTPTPTATIRHPRPYRPNVGOEALISOTTISMAPRODTSSEYIISCHPYGTDEBPLOFRVGRSTATLIG
QOOSVQOQMI FEEHGRRTPTPTATIRHPRPYRPNVGOEALISOTTISMAPRODTSSEYIISCHPYGTDEBPLOFRVGRSTATLIG
PHPGDLPNPNASVTOEALISOTTISMAPRODTSSEYIISCHPYGTDEBPLOFRVGRSTATLIG
TLTGLRGATYNIIVALKDQORHRAEYVTVGNSVNEGILNPTDSCDPYIVSHYAGD
EENRSESSEFKLCOCLGFGSGHFRCDSSRWCHNDGNVYKIGKMRQGNQGMWCT
MSTCICANGKGRKCDPHEATCYDDGKTVHVGQWQKEYGALCSCTCFGQORMRCDNGCR
NCRPRGSRPSEGTGQSYNQSORHQRNTNVNCPICEFMPLDVQADREDSRE"

RESULT 13
HSM806992 8411 bp mRNA linear HTC 23-FEB-2005
LOCUS HSM806992
DEFINITION Homo sapiens mRNA; cDNA DKFZp68601166 (from clone DKFZp68601166).
ACCESSION BX640875
VERSION BX640875.1 GI:34365169
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 8411)
BLOCKER,H., Boecker,M., Brandt,P., Mewes,H.W., Weil,B., Amid,C.,
Ossanger,A., Fobbo,G., Han,M. and Wiemann,S.
The German cDNA Consortium
Direct Submission
Submitted (22-FEB-2005) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ), Email s.wiemann@dkfz-heidelberg.de,
sequenced by GBF (National Research Centre for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.
This clone (DKFZp68601166) is available at the RZPD, Germany.
Resequencing center fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneid=DKFZp68601166
Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/.

1..8411
/gene="DKFZp68601166"
268..7701
/gene="DKFZp68601166"
/product="hypothetical protein"
/protein_id="CA845932.1"
/db_xref="GI:34365170"
/db_xref="GOA:Q6MZU5"
/db_xref="InterPro:IPR000083"
/db_xref="InterPro:IPR000562"
/db_xref="InterPro:IPR002086"
/db_xref="InterPro:IPR003961"
/db_xref="InterPro:IPR006209"
/db_xref="UniProt/TREMBL:Q6MZU5"
/translation="MLRGPGLLLAVLCIGTAVPSTGASKSRKROQWVOPSPVA
VSGSKPCYNGHNGHNOQWERTYLGNAIVCTCGSGRPNCSKREARETCDKXT
GMYRVGDTYRERKDSIMDCTCIGARGISCTIARCHEGGSYTGCTWRPHET
GGTMBECVCLANGKENTCKPIABKCTDHAAGSYVGETWERYQGMWVDCCLGE
GSGRTICTSSNRNCDOTRTSYRIGDTWSSKDKRGNLLOCITGNGMRKGCHEMTSV
QTTSSGSGPTDVAAYVQPPHPPQPPYHGSVGVSVGVGMKTOGNKMLCT
CLANGVSCFETAVLYOTGANSNEPCVLPYLYNRRYFSCITRBOGROGCTTSY
EDOKVSCFETAVLYOTGANSNEPCVLPYLYNRRYFSCITRBOGROGCTTSY
NYADOKRGCPMAAHEICTTBGWNRYRGDWDKDHGMRKCTCVNGKERTC
IAYSQLEKQICVDITVNDVTHKREBGMNLCTFEGGKGMKCDPTEQCDQDST
GTFYQIGDSWEKTVHGVRYQCYGRGIGMHCQPLCTFYPSSGPVEFPIEPSON
SHPIOMNAPQPSHISKYLIMRPNKSVGWKEATIPGLHNSYTIKGLKPEVVEGQLI
SIOOYGHOBVTRDFTTSTSTSVTVGERTPSPFLVATSVTEIIVASSVWS
VSAIDTVSGFRVVEIIEEGDEPOYDLPESTASVNPDLIPGRKYVNVYQISEDE
QSILISTQTTAPDAPDPVTDVDTISIVRRSRPOAFTTGRITYSVSBSSTIQ
NLPEIANSVTLSDLOQPOVNYITIVAEENQSTPVTIQETRTGTPSDVPSRDLQ
FVEVTVKTIWMTPEASATGVRVDIPNLGEGHQRPIRMTPEATVTSRGVT
YFKFVAVSHGRESKPLTAQOTKLDLPNLQVNESTVLRMTPEATVTSRGVT
VGLTRGQPOVNVVPSVSKYPLRNPLOPABEYVSVIYIKNGQSPKATGPFVTLQ
SSIPRYNTVETTYITVTPARIGKGVRSQGEARRETVTSBSGSIYVSGLTG
VEYVYITQVLRDQERDAPVNRVVTPLSPNULHEANPDVGLTVSWKSTPDT
GYRIITTTNGQOQNSLEEVQADQSCSTDNISPGLNEVSVYVADDESVISDT
IIEVPQLTDLSEFVDTDSIGRMTPLNSSTIIGYRIIVAAEGEIPFEDVDSV
GYVTVGLBEGIDYDISVTLINGSEAPPTLQOAVPPTDLPFNIGDPNRVTM
APPESIDITNELIVRSPVKNEDVAELISPSMNAVTLTLRGTVEYVSVSVYBOH
ESPLRGRQGTGDSPTGIDSPITANSPTVHMIAPRATITTRIRHHPHSGRPRE
DRVPHSRNSTLTNLTPGTEYVSVIALNREBSPPLIQOSTVSDVDRLEVAAP
TSLISMDAPAVTVRYRITGEGTGSPOBETVPSGKSTATISGLKPEVDTYITV
AYVGRGSPASPKRSINRYTEIDKPSOMTVQONSISKMLPSSPVTGVRVYT
PKNGPQTKTKTAGVTPOTEMTIBLOPVTYVSVYVAAONSSEGSQVVOVATNIBP
KGLAFPDVVDVSIKIMBSPOGOVSRRYVSSPBGIEHLPRAPGEERTALQGR
PGSEIYVSVVVALHDMESOPLTGTSQVSLAPYDLKTOVTVLSAQMPVQVLG
YRVATVPKERTGPMKEINLAPDSSVSATLPAKXEVSVALKOTLTSRPAQVVT
TLENVSPPRARVTDATETITISWRTKSTISCHPYGTDEBPLOFRVGRSTATLIG
YITGLOPDTYKIYUUTLNDNRSSPVVIDASTALIDAPSNRLATPNSLWQPPRARITGYI
PRARITGYIYUUTLNDNRSSPVVIDASTALIDAPSNRLATPNSLWQPPRARITGYI
PLIGRKTDELPLVTLPHNPLPILNDPSTVOKTPEVTHPDIYDNGIOLPDTSGOOSVGOOMI
FEEHGRRTPTPTATIRHPRPYRPNVGOEALISOTTISMAPRODTSSEYIISCHPYGTDEBPLOFRVGRSTATLIG
QOOSVQOQMI FEEHGRRTPTPTATIRHPRPYRPNVGOEALISOTTISMAPRODTSSEYIISCHPYGTDEBPLOFRVGRSTATLIG
PHPGDLPNPNASVTOEALISOTTISMAPRODTSSEYIISCHPYGTDEBPLOFRVGRSTATLIG
TLTGLRGATYNIIVALKDQORHRAEYVTVGNSVNEGILNPTDSCDPYIVSHYAGD
EENRSESSEFKLCOCLGFGSGHFRCDSSRWCHNDGNVYKIGKMRQGNQGMWCT
MSTCICANGKGRKCDPHEATCYDDGKTVHVGQWQKEYGALCSCTCFGQORMRCDNGCR
NCRPRGSRPSEGTGQSYNQSORHQRNTNVNCPICEFMPLDVQADREDSRE"

ORIGIN

Query Match 15.7%; Score 330; DB 4; Length 8411;
Best Local Similarity 100.0%; Pred. No. 2.1e-83;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1..8411
/organism="Homo sapiens"

OY		1764	CCCATAGCTGAGAAGGTGTTTATCATGTCTGTGGACATTCTTATGTGTGGAAAGC	1823
OY		1764	CCCATAGCTGAGAAGGTGTTTATCATGTCTGTGGACATTCTTATGTGTGGAAAGC	1823
Db		808	CCCATAGCTGAGAAGGTGTTTATCATGTCTGTGGACATTCTTATGTGTGGAAAGC	867
OY		1824	TGGAGAAAGCCCTTACCMAAGCTGATGTGTAGATTGTACTTGCTCGGAGAAAGCAGC	1883
Db		868	TGGAGAAAGCCCTTACCMAAGCTGATGTGTAGATTGTACTTGCTCGGAGAAAGCAGC	927
OY		1884	GGAAGCATTCATTGCATTCTTGAAATATGATGCCAAGATCAGAGACAAGAACATCTTAT	1943
Db		928	GGAAGCATTCATTGCATTCTTGAAATATGATGCCAAGATCAGAGACAAGAACATCTTAT	987
OY		1944	AGAAATTGAGACACCTGGAGCAAGAAAGTATA TGAGAAAACCTGCTCAGTGACATTCGC	2003
Db		968	AGAAATTGAGACACCTGGAGCAAGAAAGTATA TGAGAAAACCTGCTCAGTGACATTCGC	1047
OY		2004	ACAGCAACCGGCCGAGAGAGTAGTAGTGTGAGAGCACACCTCTGTGCAGACCAATCG	2063
Db		1048	ACAGCAACCGGCCGAGAGAGTAGTAGTGTGAGAGCACACCTCTGTGTGACACCAATCG	1107
OY		2064	AGCGGATCTGGCCCCCTTACCGGATTTGCT	2093
Db		1108	AGCGGATCTGGCCCCCTTACCGGATTTGCT	1137
RESULT 14				
LOCUS	A1095589/c			
DEFINITION	A1095589	465 bp	mRNA	linear EST 05-OCT-1998
	gb24a08.xl Soares pregnant uterus NbHPV Homo sapiens cDNA clone			
	IMAGE:1697174.3, similar to gb:X02761_cds1 FIBRONECTIN PRECURSOR			
	(HUMAN); mRNA sequence.			
ACCESSION	A1095589			
VERSION	A1095589.1	GI:3434565		
KEYWORDS	EST;			
SOURCE	Homo sapiens			
ORGANISM	Homo sapiens (human)			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;			
	Homnidae; Homo.			
REFERENCE	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.			
AUTHORS	NCI-CGAP			
TITLE	Unpublished (1997)			
JOURNAL	Contact: Robert Strausberg, Ph.D.			
COMMENT	Email: cs9apbs-r@mail.nih.gov			
	This clone is available royally-free through LNL; contact the			
	IMAGE Consortium (info@image.lnl.gov) for further information.			
	Insert Length: 1871 Std Error: 0.00			
	Seq primer: -40m13 fwd. Et from Amersham			
	High quality sequence stop: 446.			
FEATURES	Location/Qualifiers			
source	1..465			
	/organism="Homo sapiens"			
	/mol_type="mRNA"			
	/db_xref="taxon:9606"			
	/clone="IMAGB:1697174"			
	/sex="female"			
	/dev_stage="adult"			
	/lab_host="DH10B"			
	/clone_lib="Soares_pregnant_uterus_NbHPV"			
	/note="Organ: uterus; Vector: pTV73-Pac; Site 1: Not I;			
	Site 2: Eco RI; 1st strand cDNA was primed with a Not I -			
	oligo(dT) primer [5'			
	AAC TGAAGAAGATCGCGCGCCCTTTTTTTTTTTTTTTTTTTT 3'];			
	double-stranded cDNA was ligated to Eco RI adaptors			
	(Pharmacia), digested with Not I and cloned into the Not I			
	and Eco RI sites of the modified pTV73 vector. Library			
	went through one round of normalization. Library			
	constructed by M. Fatima Bonaldo."			
ORIGIN				
Query Match	15.7%; Score 329; DB 1; Length 465;			

Best Local Similarity 99.7%; Pred. No. 1,66-83;									
Matches 329; Conservative 0; Mismatches 1; Indels 0; Gaps 0;									
Qy	1754	CCCATAGCTGAGAAAGCTGTTTGGATCATGCTGTCGGACCTTCTATGTGTCGGAAACG	1823						
Db	434	CCCATAGCTGAGAAAGCTGTTTGGATCATGCTGTCGGACCTTCTATGTGTCGGAAACG	375						
Qy	1824	TGGGAGAAAGCCCTACCAAGGCTGCATGATGTGATTTGTACTTGGCTTGGAGAAAGGCAC	1883						
Db	374	TGGGAGAAAGCCCTACCAAGGCTGCATGATGTGATTTGTACTTGGCTTGGAGAAAGGCAC	315						
Qy	1884	GGAGCGATCATCTGTCATCTTGAATAATGATGACACATGACGACACAAAGACATCCAT	1943						
Db	314	GGAGCGATCATCTGTCATCTTGAATAATGATGACACATGACGACACAAAGACATCCAT	255						
Qy	1944	AGAAATGAGACACCTTGAGAGCAAGAGATATCGAGGAAACTGCTCAGTGATCTGC	2003						
Db	254	AGAAATGAGACACCTTGAGAGCAAGAGATATCGAGGAAACTGCTCAGTGATCTGC	195						
Qy	2004	ACAGGCACACGCGCCGAGAGAGTGTGAGAGGACACACTCTGTGCAGACCATCG	2063						
Db	194	ACAGGCACACGCGCCGAGAGAGTGTGAGAGGACACACTCTGTGCAGACCATCG	135						
Qy	2064	AGCGGATCTGGCCCCCTTCACCGATGTTCGT	2093						
Db	134	AGCGGATCTGGCCCCCTTCACCGATGTTCGT	105						
RESULT 15									
LOCUS	AU140802		734 bp	mRNA	linear	EST 05-AUG-2002			
DEFINITION	AU140802 PLACE4 Homo sapiens cDNA clone PLACE4000266 5', mRNA								
sequence.	AU140802								
ACCESSION	AU140802								
VERSION	AU140802.1		GI:11002323						
KEYWORDS	EST.								
SOURCE	Homo sapiens (human)								
ORGANISM	Homo sapiens								
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;								
	Hominoidea; Homo.								
REFERENCE	1 (bases 1 to 734)								
AUTHORS	Ota,T., Suzuki,Y., Saito,K., Ishii,S., Yamamoto,U., Sugiyama,T.,								
TITLE	Nishikawa,T., Nakamura,Y., Sugano,S., Masuho,Y. and Isogai,T.								
	HRI human cDNA project (Ota,T., Suzuki,Y., Saito,K., Ishii,S.,								
	Yamamoto,J., Sugiyama,T., Nishikawa,T., Nakamura,Y., Sugano,S.,								
	Masuho,Y., Isogai,T.)								
JOURNAL	Unpublished (2000)								
COMMENT	Contact: Takao Isogai								
	Genomics Laboratory								
	Helix Research Institute								
	1532-3 Yana, Kisarazu, Chiba 292-0812, Japan								
	Tel: 81-438-52-3975								
	Fax: 81-438-52-3986								
	Email: genomcs@hri.co.jp								
	HRI Human cDNA Project, 5'- & 3'-end one pass sequencing: Helix								
	Research Institute; cDNA library construction: Department of								
	Virology, Institute of Medical Science, University of Tokyo, and								
	Helix Research Institute.								
FEATURES	Location/Qualifiers								
source	1..734								
	/organism="Homo sapiens"								
	/mol_type="mRNA"								
	/db_xref="taxon:9606"								
	/clone="PLACE4000266"								
	/issue_type="Placenta"								
	/clone_id="PLACE4"								
	/note="Vector: PME185FL3"								
ORIGIN									
Query Match	15.7%;	Score 329;	DB 1;	Length 734;					
Best Local Similarity	99.7%;	Pred. No. 1,96-83;							
Matches 329; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;					

```

QY 1764 CCCATAGCTGAGAAGTGTGTTGATCATGCTGTGGGACTTCTATGTGTGAGAAACG 1823
    |||||||
Db 309 CCCATAGCTGAGAAGTGTGTTGATCATGCTGTGGGACTTCTATGTGTGAGAAACG 368
    |||||||
QY 1824 TGGGAGAAAGCCCTTACCAAGCTGATGATGTAGATTGTACTTGCTGGGAGAAAGCAGC 1883
    |||||||
Db 369 TGGGAGAAAGCCCTTACCAAGCTGATGATGTAGATTGTACTTGCTGGGAGAAAGCAGC 428
    |||||||
QY 1884 GGAAGCATCATCTTGCACTTCTAGAAATGATGCAAGATCAGAGACAAAGACATCCTAT 1943
    |||||||
Db 429 GGAAGCATCATCTTGCACTTCTAGAAATGATGCAAGATCAGAGACAAAGACATCCTAT 488
    |||||||
QY 1944 AGAATTGAGAGACACTGTGAGCAAGAAAGATATCGAGAAACCTGTCCAGTGCATCTGC 2003
    |||||||
Db 489 AGAATTGAGAGACACTGTGAGCAAGAAAGATATCGAGAAACCTGTCCAGTGCATCTGC 548
    |||||||
QY 2004 ACAAGCAAGGCGCGAGAGAGTGAAGTGTGAAGGCAACCTGTGTGACAGACCAATCG 2063
    |||||||
Db 549 ACAAGCAAGGCGCGAGAGAGTGAAGTGTGAAGGCAACCTGTGTGACAGACCAATCG 608
    |||||||
QY 2064 AGCGGATCTGGCCCTTCAACCGATGTTGGT 2093
    |||||||
Db 609 AGCGGATCTGGCCCTTCAACCGATGTTGGT 638
    |||||||

```

Search completed: February 1, 2006, 12:22:51
 Job time : 8612.84 secs

PD 27-JUN-1991.
 XX
 PF 07-DEC-1989; 89GB-00027722.
 XX
 PR 07-DEC-1989; 89GB-00027722.
 PR 07-DEC-1990; 90MO-GB001911.
 XX
 PA (BRBI-) BRITISH BIO-TECHNOLOGY LTD.
 XX
 PI Dawson KM, Hunter MG, Czapliewsk LG;
 DR MPI: 1991-208151/28.
 DR N-PSDB; AAQ02162.
 PT Fusion protein cleavage by blood clotting enzyme - for prodn. of
 PT fractions having greater antithrombotic activity for therapy and
 PT prophylaxis.
 XX
 PS Disclosure; Page 96; 115pp; English.
 XX
 CC The protein is a recombinant product of a gene fusion construct. The
 CC sequence of the synthetic hirudin HV-1 genes was designed based on the
 CC published amino acid sequence (Dott J., et al FEBS Letters 165:180
 CC (1984)). The sequence of streptokinase was obtd. from PCR amplified
 CC chromosomal DNA from S. equisimilis ATCC 10009 or ATCC 9642. The primers
 CC used for the PCR were based on the pub- lished DNA sequence of S.
 CC equisimilis strain H46A (Malke, H., Roe, B., and Ferrerelli, J.J., Gene 34
 CC 357-362 [1985]). The two sequences were used to construct an expression
 CC vector in which the hirudin gene is linked to the streptokinase gene via
 CC a linking sequence encoding a cleavage site for factor Xa.
 CC is present at the site of the target thrombus so the active agents are
 CC released specifically at the place where clot formation is occurring. See
 CC also AARI2887-R12889, AARI2891-R12894 and AARI2522. (Updated on 25-MAR-
 CC 2003 to correct PA field.)
 XX
 SQ Sequence 483 AA;

Alignment Scores:

Pred. No.: 4,24e-172 Length: 483
 Score: 1991.00 Matches: 403
 Percent Similarity: 84.29% Conservative: 10
 Best Local Similarity: 82.24% Mismatches: 27
 Query Match: 53.44% Indels: 50
 DB: 2 Gaps: 9

US-09-940-235-12 (1-2096) x AARI2885 (1-483)

QY 234 ACAGATTGACCATGCTGAGAAAGTGTTCATGCTGCGGACTTCATGTGCT 353
 |||||
 DB 4 ThrAspCysThrGlnSerGlyGln-----AsnLeuCys-----LeuCys--- 16
 QY 334 CGGAGAAACGTGGGAGAGCGGACGCGATCATCTTGCACTTGAATAATGATGCAAC 413
 |||||
 DB 17 -----GlnGlySerAsn-----ValCysGlyGlnGlyAsnLysCys--- 28
 QY 414 GATCAGACACAAGACATCTATAGATTGAGACCTCGAGCAAGAAGATATCGA 473
 |||||
 DB 29 -----IleLeuGlySerAspGlyGlnLysAsn----- 37
 QY 474 GGAACCTGCTCCAGTCATCTGCAACGCGGACGAGAGAGTGAAGTGTGAGAG 533
 |||||
 DB 38 -----GlnCysVal---ThrGlyGlnGly----- 44
 QY 534 CACACCTCTGTGCAACACATCGACCGGATCTGGCCCTTCACCGATGTT----- 584
 |||||
 DB 45 -----ThrProLysProGlnSerHisAsnAspGlyAspPheGlnLysIleProGlnGlu 62
 QY 555 -----CGTATTCGCTGCACTGAGTGGCTGTGTAACCGTCCATCTGTC 626
 |||||
 DB 63 TyrLeuGlnIleGlnGlyArgIleAlaGlyProGlnTyrIleLeuAspArgProSerVal 82
 QY 627 AACACAGCGCAATGGTGTGACGCTGCTGCTGACTGTTGAGGGGACGAATCAAGACTT 686
 |||||

DB 83 AsnAsnSerGlnLeuValValSerValAlaGlyThrValGlnGlyThrAsnGlnAspIle 102
 QY 687 AGCTTAATAATTTTGAATCGATCTAATCAACGACCTGCTCATGAGAAAGCAGAG 746
 |||||
 DB 103 SerLeuLysPhePheGlnIleAspLeuHisSerGlyProAlaHisGlyGlyLysThrGlu 122
 QY 747 CAAGGCTTAAGTCCAAATCAAAAACATTTGCTATGATGCGCGCATGTCATGAAA 806
 |||||
 DB 123 GlnGlyLeuSerProLysSerLysPhePheAlaThrAspSerGlyAlaMetProHisLys 142
 QY 807 CTTGAGAAAGCTGACTTATCAAGGCTTATCAAGAACATGATGATGCTCAACGTCACGT 866
 |||||
 DB 143 LeuGlnLysValAspLeuLeuLysAlaIleGlnGlnGlnIleAlaAsnValHisSer 162
 QY 867 AACGACGACTTATGAGGCTATGATTTTGCAGCGGATGCAACATTAAGTCAAGCAAC 926
 |||||
 DB 163 AsnAspAspLysPheGlnValIleAspPheAlaSerAspAlaThrIleThrAspArgAsn 182
 QY 927 GGCAGGCTCTACTTTGCTGACCAAGATGCTCGGTAACCTTGCCGACCCACCTGTCCA 986
 |||||
 DB 183 GlyLysValTyrPheAlaAspLysAspGlySerValThrLeuProThrGlnProValGln 202
 QY 987 GAATTTTGTAGCGGACATGTCGCGGTTAGACCATATATAAAGAAACCAATACAAAC 1046
 |||||
 DB 203 GluPheLeuLeuSerGlyHisValArgValArgProTyrLysGlnLysProIleGlnAsn 222
 QY 1047 CAAGCGAAATCTGTGATGAGATATATGATATGATGATGATGATGATGATGATGATGAT 1106
 |||||
 DB 223 GlnAlaLysSerValAspValGlnTyrThrValGlnPheThrProLeuAspProAspAsp 242
 QY 1107 GATTTGACACGAGGCTCAAGATATCAAGTATGATGATGATGATGATGATGATGATGATGAT 1166
 |||||
 DB 243 AspPheArgProGlyLeuLysAspThrLysLeuLeuLysThrLeuAlaIleGlyAspThr 262
 QY 1167 ATCAATCTCAAGATTAATTAAGTCAAGCAACAAAGATTTTAAACAAACCAACCAAGGC 1226
 |||||
 DB 263 IleThrSerGlnGlnLeuLeuAlaGlnAlaGlnSerIleLeuAsnLysThrHisProGly 282
 QY 1227 TATACGATTTATGAAGCTGACTCTCTCAATGCTCATGACATGACATTTTCCGTACG 1286
 |||||
 DB 283 TyrThrIleLysGlnArgAspSerSerIleValThrHisAspAsnAspIlePheArgThr 302
 QY 1287 ATTTTACCAATGATCAAGAGTTTACTTCAACCGTGTAAATGCGGAAACCAAGCTTAAGG 1346
 |||||
 DB 303 IleLeuProMetCaspGlnGlnLysPheThrThrHisValLysAsnArgGlnAlaLysGln 322
 QY 1347 ATCAATAAAAAATCTGCTGTAATGAAGAATAAACAACATGACCTGATCTGAGAAA 1406
 |||||
 DB 323 IleAsnLysLysSerGlyLeuAsnGlnGlnIleAsnAsnThrAspLeuIleSerGlnLys 342
 QY 1407 TATTTACGCTCTTAAAAAGGGAAGACCGTATGATCCCTTTGATGCGAGTCACTTAAA 1466
 |||||
 DB 343 TyrTyrValLeuLysLysGlyGlnLysProTyrAspProPheAspArgSerHisLysLys 362
 QY 1467 CTGTTTCAACCAATTCGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1526
 |||||
 DB 363 LeuPheThrIleLysTyrValAspValAsnThrAsnGlnLeuLeuLysSerGlnGlnLeu 382
 QY 1527 TTAACAGCTAGCGAACTGATCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 1586
 |||||
 DB 383 LeuThrAlaSerGlnAsnLeuAsnLeuAspPheArgAspLeuTyrAspProAlaGlyAspVal 402
 QY 1587 AAACCTACTTACAAACAACTCTGATCTTTTGGATTATGATGATGATGATGATGATGATGAT 1646
 |||||
 DB 403 LysLeuLeuTyrAsnAsnLeuAspAlaPheGlyIleMetCaspTyrThrLeuThrGlnLys 422
 QY 1647 GTAGAGATATATACCAATGACCAACCAACCGTATGATTAACGCTTATATAGGCGAAGCAAC 1706
 |||||
 DB 423 ValGlnAspAsnHisAspAspThrAsnArgIleIleThrValTyrMetGlyLysArgPro 442
 QY 1707 GAAGAGAGAAATGCTAGCTACATTTAGCT 1736
 |||||
 DB 443 GlnGlyGlnAsnAlaSerTyrHisLeuAla 452

RESULT 3
ADL92189
ID ADL92189 standard; protein; 413 AA.
XX
AC ADL92189;
XX
DT 20-MAY-2004 (first entry)
XX
DE Streptokinase protein sequence.
XX
KW harvesting; recombinant; host cell; N-terminal leader peptide;
KW pre-peptide; antibiotic; post-translational modification;
KW pharmaceuticals; vaccine; immunogenic.
XX
OS Unidentified.
PN WO2003099862-A1.
PD 04-DEC-2003.
XX
PF 26-MAY-2003; 2003WO-NL000389.
XX
PR 24-MAY-2002; 2002EP-00077060.
PR 07-FEB-2003; 2003US-00360101.
XX
PA (NANO-) APPLIED NANOSYSTEMS BV.
PI Moll GN, Leenhouts CJ, Kuipers OP, Driessen AJM;
DR WPI; 2004-042770/04.
PT Harvesting a desired polypeptide produced by a recombinant host cell, for
PT producing pharmaceuticals, comprises selecting a recombinant nucleic acid
PT comprising nucleic acid fragments encoding a leader peptide and the
PT polypeptide.
XX
PS Claim 4; Page 82-83; 109pp; English.
XX
CC The invention relates to a novel method for harvesting a (poly) peptide
CC produced by a recombinant host cell. The novel method involves selecting
CC a cell comprising a first nucleic acid encoding a leader peptide and a
CC second nucleic acid fragment encoding the desired (poly)peptide. The
CC first and second fragments are within the same open reading frame of the
CC first nucleic acid and the leader peptide is functionally equivalent to
CC an N-terminal leader peptide found with the pre-peptide of a lantibiotic.
CC The host cells and nucleic acids are useful for producing, harvesting and
CC post-translational modification of polypeptides. The polypeptides may be
CC used in the production of pharmaceuticals, e.g. as antigen for vaccine or
CC immunogenic composition. This sequence represents a polypeptide relating
CC to the novel method of the invention.
XX
SQ Sequence 413 AA;
XX
Alignment Scores:
Pred. No.: 1 72e-171 Length: 413
Score: 1384.00 Matches: 383
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 53.25% Indels: 0
DB: 8 Gaps: 0
US-09-940-235-12 (1-2096) x ADL92189 (1-413)
QY 588 ATTCGTGACCTGATGCTGTAGACCTCCATCTGTCAACAAGCCAAATTGTTGTT 647
Db 1 IleaIaGIyProGIuItrIpleuIeuaSpaRgProSeValaShnshnserGIuIeVal 20
QY 648 AGCGTGTCTGCTGCTGTGAGGGAGCAATCAAGACATTAGTCTTAATTTTGAATC 707
Db 21 SerValaIaGIyThrValGIuGIyThraSngIaSpIleSerIeuIySPhPeGIuIle 40
QY 708 GATCAATCATCAAGACCTGCTATGAGAGAAAGACAGACAGGCTTAATCCAAATCA 767

Db 41 AspleuThrSerIaRgProAlaIaGIyGIyThrGIuGIuGIuIeUSeRProIySeR 60
QY 768 AAACCATTTGCTACTGATAGTGGCGATGCACTATAAAGCTTGAAGCACTACTACTA 827
Db 61 LysProhealathraSpSerGIyAlaMeSerhIshIySeuGIuIyAlaIaSpIeUeu 80
QY 828 AAGGCTATTCAAGAAACATTGATGCTTAAACCTGACAGTAACGACTACTTGAAGGTC 887
Db 81 LysAlaIleGIuGIuGIuIeUleUAlaIaShnValhIshSerhshnSpaPtyrPheGIuVal 100
QY 888 ATGATTTTGCAGACGATGCAACCATTAATGATGCAACCGCAAGTCTACTTGTGAC 947
Db 101 IleSphealaseRspAlaIaThrIleThraSpRgSngIyIyValIyPheAlaSp 120
QY 948 AAGATGCTTCGTGATCTTCCGACCCCAACCTGTCCAAAGATTTTGTCTAAGCGAAT 1007
Db 121 LysSpGIySerValThrIeUProThrGIuProValGIuGIuIeUSeRGIyAla 140
QY 1008 GTGCGGTTAGACCATATTAAGAAACCAATACAAACCAAGGAAATCTGTGATG 1067
Db 141 ValhRgValaIaRgProIyRyGIuIySpRolIeGIuSngIuAlaIySerValaSpAl 160
QY 1068 GAATATACGTGATGATTAATCTCCCTTAAACCTGATGACGATTTCAACGAGTCTCAA 1127
Db 161 GIuIyThrValGIuIeUProIeUshnProhSpaSpaPheRgProGIyIeUlyS 180
QY 1128 GATCACTAAGCTATTGAAACACTAGCTACTCGGTGACACCATCACTCAAGATTTACTA 1187
Db 181 AspThryIySeuIeUlysthrIeUAlaIleGIyAspThryIeThrSergIuIeUeu 200
QY 1188 GCTCAAGCAAAAGCATTTTAAACAAACCAACCGGCTATAGATTATGAACGTGAC 1247
Db 201 AlGIuIaIaGInserIleUeUshnshIspRogIyThrIleTyGIuRgSp 220
QY 1248 TCCATCATCGTCACTGATGACAAATGATATTTCCGTGACATTTTCAAGATGATCAAG 1307
Db 221 SerSerIeValaIaThrIhshSpaShnSpIleheRgThrIleUeUProMeRAspGIuIu 240
QY 1308 TTACTATACCGTGTAAATAATCGGAAACAAGCTTATAGATCAATATAAATTCGTCTG 1367
Db 241 PheThryIyRgValIyAlaShnRgGIuGIuAlaIyRgIleShnIySergIyIeU 260
QY 1368 AATGAAGAAATTAACAACTGACCTGATCTCTGAGAAATTTTCGCTTAAATAAGG 1427
Db 261 AsnGIuIuIleShnshnThraSpIeUleSergIuIyRyIyValIeUlySergIy 280
QY 1428 GAAAGCCGATGATCCCTTGTGATCGGACGTCACTTGAACCTTCAACATCAATACGTT 1487
Db 281 GIuIyProIyIaSpProheSpaSpSerhIshIeUlySleUeUshnThrIleIyRyVal 300
QY 1488 GATGTCGATACCAAGATTTGCTTAAAGTAGAGAGCTTTTACAGCTAGCAAGTAAAC 1547
Db 301 AspValaSpThraSngIuIeUleUlySergIuGIuIeUeUThraIaSergIuRgShn 320
QY 1548 TTAGACTTCAGAGATTATACGATCTCTGATTAAGGCTAAACTTACTTACAACAATCTC 1607
Db 321 IeUaSpheRgIaSpIeUyIyAspProIaRgSpIyAlaIySleUeUlyRasShnIeU 340
QY 1608 GATGCTTTGGTATTATGAGCTATACCTTAACTGAAAGAGTAGAGATTAACAGATGAC 1667
Db 341 AspAlaIaheGIyIleUeUAspRyThrIeUThrGIySValGIuIeUAspShhIaSpSp 360
QY 1668 ACCAAGCGTATCATTAACGTTTATATAGGACCAAGGACCCGGAAGAGAGAATGCTAC 1727
Db 361 ThrAsnRgIleIeThrValIyMetGIyIySargProGIuGIuIaShnIaSerIyR 380
QY 1728 CATTTAGCT 1736
Db 381 HisIeUaIa 383
RESULT 4
AAK10194

ID AAR10194 standard; protein; 414 AA.
 XX AAR10194;
 AC
 XX
 XX 28-MAR-1991 (first entry)
 DT
 XX Strepctokinase encoded by synthetic gene.
 DE
 XX Strepctokinase; thrombolytic agent; myocardial infarction.
 KM
 XX Synthetic.
 OS
 XX EP407942-A.
 PN
 XX 16-JAN-1991.
 PD
 XX 11-JUL-1989; 89JP-00179432.
 PF
 XX 11-JUL-1989; 89JP-00179432.
 PR 27-NOV-1989; 89JP-00307957.
 PR 11-APR-1990; 90JP-00096830.
 XX
 XX (SAKA) OTSUKA PHARM FACTOR.
 PA
 XX Majima E, Ogino K, Ono K, Sakata Y, Uenoyama T;
 PI
 XX WPI; 1991-016179/03.
 DR N-PSDB; AAQ10230.
 DR
 XX Synthetic gene encoding streptokinase - scale, high purity prodn. of
 PT streptokinase used as a thrombolytic agent.
 PS
 XX Claim 1; Page 51; 76pp; English.
 XX
 CC Streptokinase and its derivatives can be produced in large quantities
 CC with high purity for use as thrombolytic agents in patients with lung
 CC thrombus or myocardial infarction. See also AAR10195-R10200
 XX
 SQ Sequence 414 AA;
 Alignment Scores:
 Pred. No.: 1.73e-171 Length: 414
 Score: 1984.00 Matches: 383
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 53.25% Indels: 0
 DB: Gaps: 0
 US-09-940-235-12 (1-2096) x AAR10194 (1-414)
 QY 588 ATTGCTGAGCTGAGTGGCTGCTAAGCCGTCATCTGTGAACAAGCCAAATTGGTTT 647
 DB 1 TlelaaglyProglutripLeuLeuAspArgProSerValAsnAspSerIleuValVal 20
 QY 648 AGCGTGTGCTGCTGCTGAGGGGAGCAATCAAGCATTAGCTTAAATTTTGGAAATC 707
 DB 21 SerValaIaaglyThrValGluGlyThraSngIaAspIleSerLeuysPheGluIle 40
 QY 708 GATCTAATCATGACGCTGCTCATGTAGAGAAAGACAGCAAGGCTTAAGTCCAAATCA 767
 DB 41 AspLeuThrSerArgProIaIaIeGlyIySthngIuGlnGlyLeuSerProIySser 60
 QY 768 AAACCATTTGCTACTGATAGTGGCGCGATGTCACATTAACCTTGAGAACTTGAATCA 827
 DB 61 LysProPhealatrAspSerGlyAlaMetSerHisLysLeuGluIlySalaAspLeu 80
 QY 828 AAGGCTATTCAAGAAATTAATGCTTAACGTCCACAGTAAGAGACTTACTTGAATC 887
 DB 81 LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAspAspIyPheGluVal 100
 QY 888 ATGATTTTGGACGATGACCACTTACTGATCGAAACGGCAAGGTCATCTTGCTGAC 947
 DB 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValIyPheAlaAsp 120

QY 948 AAGATGTTGCTGAACCTTGCCGACCCAACTGTCCAAAGATTTTGTGAAGCGACAT 1007
 DB 121 LysAspIySerValThrIleuProThrgIaProValGlnGluPheLeuLeuSerGlyHis 140
 QY 1008 GTGGCGCTTAGACCATTAAGAAACCAATCAAAACCAAGGAAATCTGTGATGTG 1067
 DB 141 ValArgValArgProIyTrLysGluLysProIleGlnSngIaIlySerValAspVal 160
 QY 1068 GAATATATCTGATGATCTTACCTCCCTTAACCCCGATGACATTTTCAAGCCAGGTCAA 1127
 DB 161 GluIyThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
 QY 1128 GATCTAAGCTAATGAAGAAACATAGTATGCGGTGACACCATCATCTCAAGAAATTA 1187
 DB 181 AspThrIyLeuLeuLysThrIleuAlaIleGlyAspThrIleThrSerGlnIleuLeu 200
 QY 1188 GCTCAAGCACAAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTATGAACGTGAC 1247
 DB 201 AlaGlnIaGlnSerIleLeuAsnLysAsnHisProGlyIyThrIleIyGluArgAsp 220
 QY 1248 TCCTCAATGCTGCTACATGACATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 1307
 DB 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
 QY 1308 TTTACTTACCGGTGTTAAATCCGGGAACAGGCTTATGAGATCAATTAATAATCTGTCG 1367
 DB 241 PheThrIyTrArgValLysAsnAspGlnGlnAlaIyTrArgIleAsnLysLysSerGlyLeu 260
 QY 1368 AATGAAGAAATTAACAACACTGACCTGATCTGTGAGAAATATTAAGCTTAAAAAGGG 1427
 DB 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysIyTrIyValIleuLysGly 280
 QY 1428 GAAAAGCCGTAATGATCCCTTGTATGCGATCACTTGAACCTGTCACCATTAATACGTT 1487
 DB 281 GluLysProIyTrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysVal 300
 QY 1488 GATGCGATACCAACGAATGTCTAAAGAGAGAGGCTTAAACAGCTAGAGAACGTGAC 1547
 DB 301 AspValAspThrAsnGluLeuLeuLysSerGlnGlnIleuLeuThrAlaSerGluArgAsn 320
 QY 1548 TTAGACTTCAGAGATTATACATCCGTCGTGATAGGCTAAAGCTTCAACATCTGC 1607
 DB 321 LeuAspPheArgAspLeuIyTrAspProIyAspLysAlaLysLeuLysIyTrAsnAsnLeu 340
 QY 1608 GATGCTTTTGGTATTAATGACATATACCTTAACCTGAAAGATAGAGATTAATCAGATGAC 1667
 DB 341 AspAlaPheGlyIleMetAspIyTrThrIleuThrGlyLysValGluAspAsnHisAspAsp 360
 QY 1728 CATTTAGCT 1736
 DB 381 HisLeuAla 383
 RESULT 5
 ID AAR63120 standard; protein; 414 AA.
 XX AAR63120;
 AC
 XX
 XX 16-OCT-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 16-NOV-1994 (first entry)
 XX
 XX Streptokinase.
 DE
 XX Streptokinase; SK; Streptococcus equisimilis; plasminogen;
 KM myocardial infarction.
 XX
 OS Streptococcus dysgalactiae subsp. equisimilis.

XX Key Location/Qualifiers
FH Region 1..352
FT /note="claim 3, see CC"
FT Region 14..414
FT /note="claim 1, see CC"
FT Region 120..352
FT /note="claim 3, see CC"
FT Region 244..414
FT /note="claim 3, see CC"
FT Region 244..352
FT /note="claim 2, see CC"
FT
FT
FN MO9407992-A1.
PN
PD 14-APR-1994.
XX
XX 05-OCT-1993; 93WO-US009502.
XX
XX 05-OCT-1992; 92US-00956692.
XX 29-SEP-1993; 93US-00128299.
XX
XX (GEHO) GEN HOSPITAL CORP.
XX (HARD) HARVARD COLLEGE.
XX
XX
XX Reed GL;
XX
XX WPI; 1994-135561/16.
XX
XX DNA encoding a polypeptide which binds to plasminogen and corresponds to
PT region of streptokinase - useful to detect plasminogen in a sample and to
PT treat myocardial infarction.
XX
XX
XX Disclosure: Page 40-41; 62pp; English.
XX
XX Nucleic acid comprising a sequence encoding amino acids 14-414 of
CC streptokinase, which binds to plasminogen and does not have a sequence
CC comprising amino acids 60-414 is new. The polypeptide pref. comprises
CC amino acids 244-352, 1-352, 120-352 or 244-414. (Updated on 25-MAR-2003
CC to correct PN field.) (Updated on 16-OCT-2003 to standardise OS field)
XX
XX Sequence 414 AA;
SQ

Alignment Scores:
Pred. No.: 1.73e-171 Length: 414
Score: 1984.00 Matches: 363
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 53.25% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-12 (1-2096) x AAR63120 (1-414)
QY 588 ATTGCTGACCTGATGGCTGCTAGACCGTCATCTGTCAACAACGCAATGTTGTT 647
Db 1 lleaaglyprrroglittrpleuenuasphaprrroserValashnshserglneuVal 20
QY 648 AGCGTGTGCTACTGTGTGAGGGGAGCAATCAAGCATTAGTCTTAAATTTTGAATC 707
Db 21 SerValaIaaglyThrValgluglyThrshenglnaaplleSerleuysphhegluile 40
QY 708 GATCTAACATCCAGACCTGCTCATGAGAGAAAGACAGACAGCGCTTAAGTCCAAATCA 767
Db 41 AspleutrrSerArprrroalshlslygllyLysrthnglulglnlyleuSerProlysser 60
QY 768 AAACCATTTGCTACTGATGATGCGGAGATGTCACTAACTTGAGAAAGCTGACTACTA 827
Db 61 LysprrrohealatrshaspsersglYalshetSerHslslyleuglulysalshaspleu 80
QY 828 AAGGCTATTCAAGAACATTTGATCGCTAACGTCACAGTAACGACGACTTATTTGAGGTC 887
Db 81 LysalatrleaglulglnleuValleashanValHisSerAsnshaspsrtyrrphegluVal 100

QY 888 ATTGATTTTTCAGACGATGCAACCATTAAGTATGAAACGCAAGGTCTACTTTGCTGAC 947
Db 101 lleaapPhehlaseraspalatrhrilleThraspashsnlslysrValYrrPhealAsp 120
QY 948 AANAGATGTTCCGTTAACTTTCGCCGACCCCAACTGCTCCAGAAATTTTGTAAAGCGACAT 1007
Db 121 LysaspGlySerValThrleuprrrothnglnprroValglnglnupheleuSerGlyHis 140
QY 1008 GTGCGGTTAGACCATTAATAAGAAAAACAATCAAAACCAAGCGAAATCGTTGATGTG 1067
Db 141 ValArgValArgProtyrlysrGluysprrroileglnasnlnalalysSerValAspVal 160
QY 1068 GAATATACGTGTACAGTTTACCTTAAACCCGTAGATGACGATTTCAAGCAGGCTCAAA 1127
Db 161 GluTyrrThrValglinPheThrProleuAsnProshaspsrPheArgProgllyleuLys 180
QY 1128 GATCTAACGCTATTGAAAAACATGACTATCGGTGACACCATCATCTCAAGAAATTA 1187
Db 181 AspThrlyleuLeuysThrleuAlaileglyAspThrilleThrserrglnglnleu 200
QY 1188 GCTCAAGCAACAAGCATTTTAAACAACCAACCGGCTATACGATTTATGAACGTCAC 1247
Db 201 AlaGlinalaGlinserrilleuasnlysnshlsprrogllyTyrrThrilleTyrgluArgAsp 220
QY 1248 TCCTCAATCGTCACTGATGACATGACATTTTCCGTACAGATTTTACCAATGATCAAG 1307
Db 221 SerSerilleValThrshaspsasnshlsPheArgThrilleuPrrroMetAspGln 240
QY 1308 TTTACTTACCGGTGTTAAAAATCGGGAACAAGCTTATAGATCAATAAAAAATCGTCTG 1367
Db 241 PheThrTyrrArgVallysrAsnArgglnglnAlaTyrrArglyleuLysSerGlyleu 260
QY 1368 AATGAAGAATAAACAACAACATGACCTGATCTCTGAGAAAAATATACGCTTAAAAAAAGG 1427
Db 261 AsnGluGluIuleasnshnThraspleuIleSerGluysTyrrTyrrValleuLysGly 280
QY 1428 GAAGAAGCCGATGATCCCTTATGCGAGTCACCTTGAAATGTTACCATCAATACGTT 1487
Db 281 GluLysProtyrAspProPheAspArgSerHisleuLysleuPrrroThrilleLysTyrrVal 300
QY 1488 GATGTCGATACCAACGAAATGTCTAAAAAGTACAGAGCTCTTAAACGCTACGAACTAAC 1547
Db 301 AspValAspThrshnshngluLeuLeuLysSerGluGlnleuThrAlaserGluArgAsn 320
QY 1548 TTAGCTTCAAGAGATTTATTCGATCTCTGCTGATPAGGCTTAACTCTTCAACAATCTC 1607
Db 321 LeuAspPheArgAspleuTyrrAspProArgAspLysAlaLysleuLeuTyrrAsnAsnleu 340
QY 1608 GATGCTTTGTGATTTATGACCTATACCTTAACTGSAAGAAATAGAGATTAATCAAGATAC 1667
Db 341 AspAlaPheGlylleMetAspTyrrThrleuThrGlyLysValGluAspAsnHisAspAsp 360
QY 1668 ACCAAGCTATCATTAACCGTTTATATGCGCAAGCAGACCCGAAGAGAGAAATGCTAGCTAC 1727
Db 361 ThrAsnArggilleleThrValTyrrMetGlyLysArgProglugluAsnAlaserTyrr 380
QY 1728 CATTAGCT 1736
Db 381 HisleuAla 383

RESULT 6
AA124794
ID AAY24794 standard; protein; 414 AA.
XX
XX AAY24794;
AC
XX
XX 17-OCT-2003 (revised)
DT 26-AUG-1999 (first entry)
XX
XX Streptococcus equisimilis native streptokinase.
DE
XX Streptococcus; streptokinase; fibrin-dependent plasminogen activator;
KW NSK; RSK; bacterial; blood clot; thrombotic condition;

KM myocardial infarction; venous thrombosis; pulmonary embolism;
 KM cerebral thrombosis; graft thrombosis; arterial thrombosis.
 XX Streptococcus dysgalactiae subsp. equisimilis.
 XX MO9931247-A1.
 XX 24-JUN-1999.
 PD 15-DEC-1998; 98WO-US026694.
 PF 15-DEC-1997; 97US-0069497P.
 XX 15-DEC-1997; 97US-0069497P.
 PR (HARD) HARVARD COLLEGE.
 PA
 XX Read Gl;
 PI WPI, 1999-395183/33.
 DR N-PSDB; AAX80492.
 PS N-terminally deleted streptokinase.
 XX
 PS Claim 30; Page 60-61, 73pp; English.
 CC The present invention describes an isolated bacterial protein that
 CC induces fibrin-dependent plasminogen activation in a pharmaceutical
 CC composition for dissolving blood clots. Also described are: (1) a
 CC composition comprising an isolated modified streptokinase, the
 CC modification being removal of amino acid residues in the amino terminus;
 CC (2) a method for dissolving a blood clot in a subject, comprising
 CC administering to the subject a fibrin-dependent streptokinase protein; a
 CC nucleic acid (1) encoding a modified bacterial streptokinase; (3) an
 CC expression vector comprising (1); and (4) a host cell transformed with
 CC the expression vector of (3). The pharmaceutical composition comprising a
 CC bacterial fibrin-dependent plasminogen activator is useful for dissolving
 CC blood clots in patients with a thrombotic condition, e.g. myocardial
 CC infarction, venous thrombosis, pulmonary embolism, cerebral thrombosis,
 CC graft thrombosis and arterial thrombosis. The modified streptokinase can
 CC also be used in non-human mammals. Streptokinase activation of
 CC plasminogen is at least 10-fold, preferably 100-fold greater in the
 CC presence of fibrin than in the absence of fibrin. The modified
 CC streptokinase has at least one amino acid substitution that inactivates a
 CC substrate site for proteolytic cleavage. This reduces the rate of
 CC degradation of the streptokinase at least two-fold. The present sequence
 CC represents native streptokinase (nsk). (Updated on 17-OCT-2003 to
 CC standardise OS field)
 CC
 XX
 XX
 SQ Sequence 414, AA;
 Alignment Scores:
 Pred. No.: 1.73e-171 Length: 414
 Score: 1984.00 Matches: 383
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 53.25% Indels: 0
 DB: 2 Gaps: 0
 US-09-940-235-12 (1-2096) x AAY24794 (1-414)
 QY 588 ATTGCTGACCTGAGTGGCTGCTAGACGCTCATCTGTCAACAAGCAATGTTGTT 647
 DB 1 TlealaglyProglutirpneuuspapgProseValashnserscIneuVal 20
 QY 648 AGCGTGTGCTGCTGTGAGGGAAGATCAAGACATTAGTCTTAATTTTGAATC 707
 DB 21 SerValalaglyThrValGluGlyThrangInaspIleSerLeuysfhepneGlu 40
 QY 708 GATCTAACATCAAGCACTGCTCATGAGAAAGACAGACAGCTTAAGTCCAAATCA 767
 DB 41 AspLeuThrSerArgProAlahIsglyGlyThrGluGlnGlyLeuSerProlySer 60
 QY 768 AAACATTGCTGACATGATGAGCGGAGTGCATTAACCTTAAGAAAGCTGACTTA 827

DB 61 LysProhealathraspSerglyalameSerhislyseuGluValaspleu 80
 QY 828 AAGCTATTCAAGAACTGATGCTTACGCTACAGCACTTACTTGAAGTCC 887
 DB 81 LysAlallegIngluInleuIlealasnValhisSerasnaspPtyrPheGluVal 100
 QY 888 ATTGATTTGCAAGCGATGCAACCATTACTGATCGAAAACGCAAGTCTTACTTGTGAC 947
 DB 101 IleaspPhealaserapIatThrIleThrAspArganglyValTyPhealasp 120
 QY 948 AAAGATGTTGCGTGAACCTTCCGACCACTGCTCCAGAAATTTTGTAGCGACAT 1007
 DB 121 LysaspGlySerValThrLeuProThrGlnProValGlnGluPheLeuSerGlyHis 140
 QY 1008 GTGCGGTTAGACCATTAAGAAACCAATACAAACCAAGGAAATCGTTGATGTCG 1067
 DB 141 ValArgValArgProTyrrLySGluLyProlIeGlnangInalalySerValaspVal 160
 QY 1068 GAATATACGTACAGTTTACTCCCTTAAACCTGATGACGATTCAGACCAAGTCTCAAA 1127
 DB 161 GluTyrrThrValGlnPheThrProLeuasnProaspaspPheArgProGlyLeuLys 180
 QY 1128 GATCTAAGCTATTGAAAACACTAGCTATGCTGACACCATCATCTCAAGAAATTA 1187
 DB 181 AspThrlyseuLeuysThrLeuAlaleGlyAspThrIleThrSergIngluLeu 200
 QY 1188 GCTCAACGCAAGACATTTTAAACAAACCAACCAAGCTTACGATTTATGAAACGTCAC 1247
 DB 201 AlaglnIleGlnserIleLeuasnLyasnHisPProGlyTyrrThrIleTyrrGluArgasp 220
 QY 1248 TCCCTCAATCGTCACTGACAAATGACATTTTCCGTGAGATTTTACCAATGATCAAGAG 1307
 DB 221 SerSerIleValIthrHisaspasnaspIlePheargThrIleLeuProMetaspGlnGlu 240
 QY 1308 TTACTTACCGGTGTTAAAAATCGGAAACAAGCTTATGATCAATTAATAATCGTCTG 1367
 DB 241 PheThrTyrrArgVallyasnAlargIngluAlatyrArgIleasnLylysSerglyLeu 260
 QY 1368 AATGAGAAATTAACAACACTGACCTGATCTCTGAGAAATTTACGTTTAAAAAGGG 1427
 DB 261 AsnGluGlnIleasnThrAspLeuIleSergIuysTyrrValleuLySlyGly 280
 QY 1428 GAAAGCGGATGATGCTTGAATGCGAGTCACTTGAACCTGTCACCAATCAATACGTT 1487
 DB 281 GluLysProTyrraspProPheaspArgSerHisLeuysLeuPheThrIleTyrrVal 300
 QY 1488 GATGCGATCAACAAGATTTGCTAAAGAGAGAGCTTTAACAGCTACGAACTTAAC 1547
 DB 301 AspValaspThrangIuleuLeuLysSergIuGlnLeuThrIleSergIuArgasn 320
 QY 1548 TTAGACTTCAGAGATTTATTCGATCCTCGTATAGGCTTAACCTTCAACATCTTC 1607
 DB 321 LeuaspPheargaspLeuTyrraspProArgaspIlyalalySleuLeuTyrrasnAsnLeu 340
 QY 1608 GATGCTTTGCTTATTTAGCTATTAACCTTAACTGAAAGATAGAGATTAATCAAGAGAC 1667
 DB 341 AspAlaspheGlylleMetaspTyrrThrLeuThrGlyValGluaspasnHisaspasp 360
 QY 1668 ACCAACGATCAATTAACCGTTTATATGAGGAGAGACCGGAAAGAGAAATGCTAGTAC 1727
 DB 361 ThrasnArgIleIleThrValItyrMetGlyLysArgProGluGluasnAlaserTyrr 380
 QY 1728 CATTTAGCT 1736
 DB 381 HisleuAla 383
 RESULT 7
 ID AAW94664 standard; protein; 414 AA.
 XX AAW94664;
 AC AAW94664;
 XX
 DT 17-OCT-2003 (revised)

DT 04-MAY-1999 (first entry)
 XX Streptococcus equisimilis native streptokinase.
 XX Streptococcus equisimilis H46A; streptokinase; mutant; fibrinolytic;
 KW plasma clot; hydrolysis; haemolytic Streptococcus; plasminogen; plasmin;
 KW serine protease; fibrin; blood clot; thrombolytic;
 KW vascular thrombolytic symptom; acute myocardial infarction;
 KW fibrinolysis; resistance.
 XX Streptococcus dysgalactiae subsp. equisimilis.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 14 /note= "encoded by ACC"
 XX
 XX US5876999-A.
 XX 02-MAR-1999.
 XX 06-DEC-1995; 95US-00568393.
 XX 06-DEC-1995; 95US-00568393.
 XX (NASC-) NAT SCI COUNCIL.
 XX Wu H;
 XX WPI; 1999-189643/16.
 DR N-PSDB; AAX16632.
 XX
 PT Mutant streptokinase polypeptide - useful as plasmin-resistant
 PT chromolytic agent.
 XX
 PS Claim 1; Col 7-10; 17pp; English.
 XX
 CC The present invention describes a mutant streptokinase (SK) polypeptide
 CC in which at least one of the amino acids in the Pro58-Lys59-Ser60-Lys61
 CC segment of the corresponding native SK is replaced by another amino acid.
 CC The present sequence represents native SK. SK is a secretory protein of
 CC haemolytic Streptococcus able to activate human plasminogen (HPIg) to
 CC plasmin (HPIa), which is a serine protease able to catalyse the
 CC hydrolysis of fibrin in blood clots. The SK is useful as a thrombolytic
 CC agent in the treatment of vascular thromboembolytic symptoms such as
 CC acute myocardial infarction. Compared with wild-type SK, the K59S mutant
 CC is more resistant to degradation by human plasmin and is more effective
 CC both in acting as a fibrinolytic agent and in activating human plasminogen.
 CC (Updated on 17-OCT-2003 to standardise OS field)
 CC
 XX
 SQ Sequence 414 AA;
 Alignment Scores:
 Pred. No.: 1,73e-171 Length: 414
 Score: 1984.00 Matches: 383
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 53.25% Indels: 0
 DB: 2 Gaps: 0
 US-09-940-235-12 (1-2096) x AA094664 (1-414)
 QY 588 ATTCGTGACCTGAGTGTCTGTAGACCGTTCATCTGTCAACAACGCAATGTTGTT 647
 Db 1 IleaIaGIyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
 QY 648 AGCGTTGCTGACTGTTGAGGGGAGCATCAAGATTAAGTTAAATTTTGAATC 707
 Db 21 SerValaIaGIyThValGIuGIyThAsnGlnAspIleSerLeuLysPhePheGluIle 40
 QY 708 GATCAATACATGACGCTGCTCATGAGAGAAAGAGAGAGAGAGGCTTAAGTCCAAATCA 767
 Db 41 AspLeuThrSerArgProIaIaHisGIyGIyThGlnGlnIleuSerProLysSer 60

QY 768 AACCATTTGCTACTGATAGTGGCGCATGTCATATAACTTGAGAAAGTGACTTACTA 827
 Db 61 LyeProPheIaIaThrAspSerGIyAlaMetSerHisLysLysLeuGIuLysValAspLeuLeu 80
 QY 828 AAGCTATTCAAGAAACATTTGATGCTTACGCTCAGAGTAAAGAGACTTGTAGAGTC 887
 Db 81 LysAlaIleGIuGlnGlnIleuIleAlaAsnValHisSerAsnAspArgTyPheGIuVal 100
 QY 888 ATTGATTTTGCAGACCATGCAACCATTACTGATGAAACGGCAAGGTCTACTTGTGCAC 947
 Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGIyLysValTyPheAlaAsp 120
 QY 948 AAGATGCTTCGGTAACTTCGCGAACCCGATCCGATCAAGAATTTTGTGAAGGACAT 1007
 Db 121 LysAspGIySerValThrLeuProThrGlnProValGlnGlnPheLeuLeuSerGIyHis 140
 QY 1008 GTGCGCTGTAGACCATTAATAAGAAACCAATACAAACCAAGGCAAAATCTGTGATGTG 1067
 Db 141 ValArgValArgProTyLysGIuLysProIleGlnAsnGlnAlaLysSerValAspVal 160
 QY 1068 GAATATACGTGTACAGTTTACTCCCTTAAACCTGATGACGATTCAGACGAGTCTCAA 1127
 Db 161 GluTyThrValGlnPheThrProLeuAsnProAspAspArgProGIyLeuLys 180
 QY 1128 GATACATAAGCTATTGAATAACACTAGCTATGCTGTGACACCATCATCTCAAGAAATTA 1187
 Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGIyAspThrIleThrSerGlnGlnLeuLeu 200
 QY 1188 GCTCAAGCAAAAGCATTTTAAACAAACCAACACGAGCTATACGATTTATGAACGTGAC 1247
 Db 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGIyTyThrIleTyGIuArgAsp 220
 QY 1248 TCTCAATGCTCACTCATGACATGACATGATTTCCGTAAGATTTTACCAATGATCAAG 1307
 Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
 QY 1308 TTTACTTACCGTGTGTTAAATTCGGGAACAGCTTATAGATCAATTAATAATTCGCTGTG 1367
 Db 241 PheThrTyThrValLysAsnArgGIuGlnAlaTyArgIleAsnLysLysSerGIyLeu 260
 QY 1368 AATGAAGAATTAACCAACACTGACCTGATCTGAGAAATATTAAGTCTTAAAAAGG 1427
 Db 261 AsnGIuGlnIleAsnAsnThrAspLeuIleSerGIuLysTyTyValLeuLysLysGIy 280
 QY 1428 GAAAGCCGTATGATCCCTTGTATGCGAGTCACTTGAACCTGTTCACCATCAATACGTT 1487
 Db 281 GluLysProTyThrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyVal 300
 QY 1488 GATGCGATACCAACGAATTCCTTAAAAAGTACAGCTCTTAAACGCTAGCGGAAGTAA 1547
 Db 301 AspValAspThrAsnGlnLeuLeuLysSerGIuGlnIleuLeuThrAlaSerGIuArgAsn 320
 QY 1548 TTAGACTTCAGAGATTATATGATCCGATGATAAGGTAAAGTACTTACCAACAATCTC 1607
 Db 321 LeuAspPheArgAspLeuTyAspProArgAspLysAlaLysLeuLeuTyAsnAsnLeu 340
 QY 1608 GATGCTTTGTGATATATGACATATACCTTATCGTGAAGAAAGTATGAGATATACGATAC 1667
 Db 341 AspAlaIaPheGIyIleMetAspTyThrLeuThrGIyLysValGIuAspAsnHisAspAsp 360
 QY 1668 ACCAACCGTATCATACCGTTTATATGAGGAGAGAGAGAGAGAGAGATATAGCTTAC 1727
 Db 361 ThrAsnArgIleIleThrValTyMetGIyLysArgProGIuGIuLysAlaLysTy 380
 QY 1728 CATTTAGCT 1736
 Db 381 HisLeuAla 383
 RESULT 8
 AA01556
 ID AA01556 standard; peptide: 414 AA.
 XX
 AC AA01556;

XX 17-OCT-2003 (revised)
 DT 18-JUN-1999 (first entry)
 XX
 DE Native streptokinase protein sequence.
 XX Antigenic peptide; streptokinase; streptokinase-specific antibody;
 KW thrombolytic activity; thrombolytic therapy; glomerulonephritis;
 KM rheumatic fever.
 XX
 OS Streptococcus dysgalactiae subsp. equisimilis.
 PN WO9908698-A1.
 PD 25-FEB-1999.
 XX
 PF 18-AUG-1998; 98WO-US017114.
 XX
 PR 18-AUG-1997; 97US-0055911P.
 XX
 PA (HARD) HARVARD COLLEGE.
 XX PA (GENO) GEN HOSPITAL CORP.
 XX
 PI Reed GL, Parhami-Seren B;
 XX
 DR WPI; 1999-190113/16.
 XX
 PT New polypeptides which bind streptokinase-specific antibodies - useful in
 PT thrombolytic therapy.
 XX
 PS Disclosure; Page 12; 44p; English.
 XX
 CC The present sequence represents a native streptokinase. The specification
 CC describes a polypeptide which binds to a streptokinase-specific antibody
 CC and prevents the antibody binding to native streptokinase. The
 CC specification also describes a synthetic polypeptide (PI) comprising an
 CC epitope which binds to a streptokinase-specific antibody and reduces
 CC thrombolytic activity of streptokinase. PI is used in thrombolytic
 CC therapy, and to prevent or treat glomerulonephritis and rheumatic fever.
 CC (Updated on 17-OCT-2003 to standardise OS field)

SQ Sequence 414 AA;

Alignment Scores:

Pred. No.: 1,73e-171 Length: 414
 Score: 1984.00 Matches: 383
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 53.25% Indels: 0
 DB: 2 Gaps: 0

US-09-940-235-12 (1-2096) x AAY01556 (1-414)

QY 588 ATTGCTGACCTGAGTGGCTGCTAGACCGTGCATCTGTCAACAGCAAGCAATTGGTTGT 647
 DB 1 IleaIaIaGlyProGIuItrPleuIeuAaPArqProSeValaAaMaSnerGIuIeuValVal 20
 QY 648 AGCGTTGCTGTAAGTGTGAGGGAGCAATCAAGACATTAGTCTTAATTTTGAATC 707
 DB 21 SerValaIaGlyThrValaGIuGIuThraaGlnAaPrlIeSerIeuIuSPrheGluIle 40
 QY 708 GATCTTAACATCAAGACCTGTCTCATGGAGAAAGACAGACAGCTTAAGTCCAAATCA 767
 DB 41 AspIeuThrSerArqProAlaHISglYglYlYThGIuGlnGIuIeuSerProIySer 60
 QY 768 AAACGATTGTACTGATGAGGGCGGAGTGCATAAATTGAGAAAGCTGACTTA 827
 DB 61 LysProPheAlaIthrAspSerGIyAlaMetSerHISLysIeuGIuIySaIaAspIeuIeu 80
 QY 828 AAGGCTATTCAAGAACAAATTGATCGCTTAACGTCCAGATTAACAGCACTATTGAGGTC 887
 DB 81 LysAlaIleGIuGIuGIuIeuIleAlaAaHISerHISaMaSnerAaPArqPrrheGluVal 100

QY 888 ATTGATTTCGACAGCATGACCAACATTAAGTCAAGAAAGCAAGTCTACTTGTGAC 947
 DB 101 IleasPheAlaSerAspAlaIthrIleThraPaaGaaGlyIyValIyrrheAlaAsp 120
 QY 948 AAAGATGTTCCGTACACCTTCGCCAGCCAACTGTCTCAAGAAATTTTGTAAAGCGACAT 1007
 DB 121 LysAspGIySerValThrIeuProThrGIuProValGIuGIuPheIeuIeuSerGIyHIS 140
 QY 1008 GTGGCGGCTTAACCATTAATAAGAAACCAATTAACAACCAAGCAAAATCTGTGATGTG 1067
 DB 141 ValaIyValaIyProIyrrLysGIuIyProIleGIuAaGlnAlaIySerValaAspVal 160
 QY 1068 GAATATACGTGACAGTTACTCCCTTAACCCCTGATGACGATTTCAGACCAAGTCTCAA 1127
 DB 161 GIuIyThrValaGIuPheThrProIeuAaSnProAaPaaPaaPaaPaaPaaPaaPaaPaa 180
 QY 1128 GATACCTAAGCTATTGAAAAACATGACTTTCGGTGAACACATCACTCAAGAAATTA 1187
 DB 181 AsprHrIySLeuIeuIyThrIeuAlaIleGIyAspHrIleThrSerGIuIeuIeu 200
 QY 1188 GCTCAAGCAACAAGCATTTTAAACAAAACCAAGCGCTTATGATTTATGACGTGAC 1247
 DB 201 AlaGlnAlaGlnSerIleIeuAaSnLysAaSnHISProGIyIyrrIleIyrrGIuIyAsp 220
 QY 1248 TCCCTCAATCCGTCACTGACATGACATGACATTTTCCGTACGATTTTAACAAATGATCAAG 1307
 DB 221 SerSerIleValIthrHISAspAaSnAspIlePheAaGlnHrIleIeuProMetAspGIuIu 240
 QY 1308 TTTACTTAACCGTGTGTTAAATCGGAAACAACTTATAGATCAATTAATAAAATCTGCTG 1367
 DB 241 PheThrIyrraIyrrValyAsaAaIyGIuGlnAlaIyrraIyrrIleAaIySLeuIySerGIyIeu 260
 QY 1368 AATGAAGAATTAACAACATGACCTGATCTCTGAAATATTAAGTCTTTAAAGAGG 1427
 DB 261 AsnGIuGIuIleAaAaSnThraSpleuIleSerGIuIyrrIyrrValIeuIySLeuGIy 280
 QY 1428 GAAAGCGATGATGATCCCTTGATCGAGTCACTTGAACCTGTTAAGTCAACATTAAGCTT 1487
 DB 281 GIuIySProIyrraPProPheAspArqSerHISLeuIySLeuPheThrIleIyrrVal 300
 QY 1488 GATGTCGATACCAACGAATGCTTAAGAAAGTGAAGAGCTTTAAGCTAGCGAAGCTAAC 1547
 DB 301 AspValaAspHraaGlnIeuIeuIySLeuIySerGIuGIuIeuIyrrAlaSerGIuIyrraAa 320
 QY 1548 TTAGACTTCAGAGATTATTAACATCTCTGTGATTAAGCTTAAGCTTACATCAACATCTC 1607
 DB 321 LeuAspPheArqAspIeuIyrraPProIyrraPAspIyrraIySLeuIyrraAaSnIeu 340
 QY 1608 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAGAAAGTAAAGATTAACAGATGAC 1667
 DB 341 AspAlaPheGIyIleMetAspIyrrThrIeuThrGIySValaGIuAspAaSnHISAspAsp 360
 QY 1668 ACCAAGCGTATCAATTAACGTTTATATGAGGAGCAAGCGAAGGAGAAATGCTAGCTAC 1727
 DB 361 ThrAaMaIyIleIeuIyrrValIyrrMetGIyLysArqProGIuGIuIeuAlaSerIyrr 380
 QY 1728 CATTAGCT 1736
 DB 381 HisIeuAla 383

RESULT 9
 AAY90282
 ID AAY90282 standard; protein; 414 AA.
 XX
 AC AAY90282;
 XX
 DT 12-SEP-2003 (revised)
 DT 13-OCT-2000 (first entry)
 XX
 DE S. equisimilis streptokinase.
 XX
 KW Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
 KW plasminogen; human; fibrinectin; thrombolytic therapy;

KM cardiovascular disorder; fibrinectin.
 XX Streptococcus dysgalactiae subsp. equisimilis.
 XX BP1024192-A2.
 XX 02-AUG-2000.
 XX 23-DEC-1999; 99EP-00310541.
 XX 24-DEC-1998; 98IN-DE003825.
 XX (COUL) CSIR COUNCIL SCI IND RES.
 XX Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
 PI Yadav M;
 XX WPI; 2000-516032/47.
 DR N-PSDB; AAA37633.
 XX Hybrid streptokinase-fibrin binding domain polypeptides useful for
 PT thrombolytic therapy comprises a streptokinase fused with fibrin binding
 PT domains of human fibrinectin.
 XX
 PS Example 3; Fig 3; 58pp; English.
 XX This sequence represents the human Streptococcus equisimilis
 CC streptokinase protein sequence. The invention relates to a hybrid
 CC plasminogen activator (PA) comprises a polypeptide fusion between
 CC streptokinase (SK), which are capable of plasminogen (PG) activation, and
 CC fibrin binding regions of human fibrinectin, which are from fibrin
 CC binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA possesses the
 CC ability to bind with fibrin independently and also characteristically
 CC retains a PG activation ability which becomes evident only after a
 CC pronounced duration, or lag, after exposure of the PA to a suitable
 CC animal or human PG. The hybrid streptokinase-fibrin binding domain
 CC polypeptides are useful in thrombolytic therapy for various kinds of
 CC cardiovascular disorders. The hybrids have enhanced fibrin selectivity as
 CC well as kinetics of plasminogen activation that are distinct from that of
 CC natural streptokinase in being characterised by a temporary delay, or lag
 CC of several minutes in the natural rate of the catalytic conversion of
 CC plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins
 CC can bind tightly with fibrin in blood clots soon after introduction into
 CC the vascular system without significantly activating the circulating
 CC blood plasminogen to plasmin, thus aiding in the localisation of the
 CC plasminogen activation process to the site of pathological thrombus. This
 CC overcomes systemic plasminogen activation encountered during clinical use
 CC of streptokinase. (Updated on 12-SEP-2003 to standardise OS field)
 CC
 XX Sequence 414 AA;
 SQ
 Alignment Scores:
 Pred. No.: 1.73e-171 Length: 414
 Score: 1984.00 Matches: 383
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 53.25% Indels: 0
 DB: 3 Gaps: 0
 US-09-940-235-12 (1-2096) x AAY90282 (1-414)
 QY 588 ATTTGCTGACCTGAGTGGCTGCTAAGCCGTCATCTGTCAACAAGCCAAATGTTGTT 647
 DB 1 TlelaaglyProglutripheuleuaspArgProSeValasnaserserGlnleuVal 20
 QY 648 AGCGTGTGAGTACGTTGAGGGGAGCATCAAGCATTAAGTCTTAATTTTGTGAATC 707
 DB 21 ServAlaaglyThValGlnGlyThrAsnGlnAspIleSerleuysPheGlnIle 40
 QY 708 GATCTAACATCAAGCACTGCTCATGAGAGAAAGAGACAGACAGGCTTAAGTCCAAATCA 767
 DB 41 AspleuthrserArgProAlahisglYgIystrGlnGlnGlyleuSerProlysser 60

QY 768 AAACATTTGCTACGTAGTGGCGGAGATGTCAATTAACCTTGAGAAAGCTGACTTA 827
 DB 61 lyePhePheAlathrAspSerGlyAlaMetSerHislylsleuGlnlyValaAspLeu 80
 QY 828 AAGCTAATTCAGAACAAATTTGATCGCTAACGTCACAGTAACGACTTGTGAGTTC 887
 DB 81 lyeAlaIleGlnGlnleuIleAlaAsnValHisSerAsnAspArgTyrPheGlnVal 100
 QY 888 ATTTGATTTTGCAGGCGATGCAACCATTAATTCGATCGAAACGGCAAGGCTACTTTCGAC 947
 DB 101 IleAspPheAlaSerAspAlathrIleThrAspArgAsnGlylyValTyrPheAlaAsp 120
 QY 948 AAAGATGTTGGTAACTTTCGCGACCAACCTGCTCAAGAAATTTTGTAAAGCGACAT 1007
 DB 121 lyeAspGlySerValThrleuProthGlnProValGlnIleuPheleuSerGlyHis 140
 QY 1008 GTGGCGGTTAGACCATTAATAAGAAAAACAATAAACAAGCGAAATCTTTGATGTG 1067
 DB 141 ValArgValArgProTyrlysgIuysProIleGlnAsnGlnAlaIysSerValAspVal 160
 QY 1068 GAATTAACCTGTACAGTTTACTCCCTTAAACCTCGATGACGATTTTCAGACCGCTTCAA 1127
 DB 161 GluTyrThrValGlnPheThrProleuAsnProAspAspArgProGlyleuLys 180
 QY 1128 GATACCTAAGCTATTGAAAACTAGCTATCGTGACACCATTCATCTCAAGAAATTTACTA 1187
 DB 181 AspThrlylsleuLeuylstrIleuAlaIleGlyAspThrIleThrserGlnIleuLeu 200
 QY 1188 GCTCAAGCAACAAAGATTTTAAACAAACAAACACCGAGCTATACGATTTTGAACGTGAC 1247
 DB 201 AlaGlnAlaGlnserIleleuAsnlyAsnHisProGlyTyrThrIleTyrGlnIuArgAsp 220
 QY 1248 TCTCTCATGTCTACTATGACATGACATTTTCCGATCGAATTTTCCAAAGATCAAGAG 1307
 DB 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleuProMetAspGlnGln 240
 QY 1308 TTACTCTACCGGTAAAAATGGGAAACAAGCTTATAGATCAATAAAAATCGTGGCTG 1367
 DB 241 PheThrTyrArgValIlylsAsnArgGlnGlnAlaTyrArgIleAsnlySlySseGlyLeu 260
 QY 1368 AATGAAGAAATAAACAACACTGACTGTCTGAGAAATATTACTGCTTAAAAAAAGG 1427
 DB 261 AsnGlnGlnIleAsnAsnThrAspLeuIleSerGlnlystrTyrValleuIlylyGly 280
 QY 1428 GAAACCCGTATGATCCCTTTGATGCGAGTCACTTGAACSTGTTACCATCAATACGTT 1487
 DB 281 GluTyrProTyrAspProPheAspArgSerHisleuIysleuPheThrIleIystrVal 300
 QY 1488 GATGTGATACCAACGAATTTGCTAAAGAGACAGCTTACAGCTAGCGAAGCGTAAC 1547
 DB 301 AspValAspThrAsnGlnleuLeuIysSerGlnGlnleuLeuThrAlaSerGlnArgAsn 320
 QY 1548 TTAGCTCGAGAGATTATATACATCCCTCGATAGAGCTAAACTACTCAACAAATCTC 1607
 DB 321 leuAspPheArgAspIleuTyrAspProArgAspIlyAlaIysleuLeuTyrAsnAsnleu 340
 QY 1608 GATGCTTTTGGTATTAAGCATATACCTTAACSTGAAAGTAAGAGATAATCAGATGAC 1667
 DB 341 AspAlaPheGlyIleMetAspTyrThrIleuThrGlyIysValGlnAspAsnHisAspAsp 360
 QY 1668 ACCAACCCTATCATTAACCGTTTATATGGGAGACGACCCGAAGAGAAATGCTAGTAC 1727
 DB 361 ThrAsnArgIleIleThrValTyrMetGlyIysArgProGlnGlnIleuAlaSerTyr 380
 QY 1728 CATTTAGCT 1736
 DB 381 HisleuAla 383
 RESULT 10
 ID AAR11829 standard; protein; 531 AA.
 AC AAR11829;

[illegible]

D	b	197	LeuYsaIaIleGInGInGInLeuIleIleAlaSnVaIhISeSerAsnAspAspTYrPheGlu	216
Q	y	885	GTCAATGATATTTGGCAAGCGATGCAACCAATTACTGATCGAAACGGCAAGGTCTACTTTGCT	944
D	b	217	ValIleAspPheAlaSerAspAlaThrIleThrAspArgSngInGlyValTYrPheAla	236
Q	y	945	GACAAAGATGGTTGGGTATACCTTGGCCGACCCAACTGTCCCAAGAAATTTTGGTAAAGCGGA	1004
D	b	237	AspLYsaSPGlySerValIThrLeuProthGInProValGInGInPheLeuLeuSerGly	256
Q	y	1005	CATGTGCGCGTTAAACCATATTAAGAAAGAAACCAATACAAAGCAAGCGAAATCGTTGAT	1064
D	b	257	HisValrArgValaArgProTYrGlySgInLYrProIleGInSngInAlaLYrSerValAsp	276
Q	y	1065	GTGGATATATACCTGACAGTTTACTCCCTTAAACCCGTGATGACGATTTTCAGACCAAGCTCTC	1124
D	b	277	ValGInTYrThrValaGInPheThrProLeuAsnProAspAspPheAlaArgProGlyLeu	296
Q	y	1125	AAAGATATCTAAGCATTTGAAAAACACTAGCTATCGGTGACACCATCAATCTCAAGAAATTA	1184
D	b	297	LYsaPheThrLYrLeuLeuLYrThrLeuAlaIleGlyAspThrIleThrSerGInGInLeu	316
Q	y	1185	CTACTCAAGACCAAGACATTTTAAAACAAAAAACACCCAGCGCTTACGATTTTATGAAAGCT	1244
D	b	317	LeuAlaGInAlaGInSerIleLeuAsnLYsaSnhIleProGlyTYrThrIleTYrGInLYr	336
Q	y	1245	GACTCTCAATCGTCACTGATGACAAATGACATTTTCCGTACGATTTTACCAATGATCAAA	1304
D	b	337	AspSerSerIleValIThrhIAspAsnAspIlePheArgThrIleLeuProMetAspGln	356
Q	y	1305	GAGTTTACTTACCGTGTAAAAATCGGGAACAGCTTATAGATCAATATAAAAAATCTGGT	1364
D	b	357	GluPheThrLYrArgValaLYrAsnArgGInGInAlaTYrArgIleAsnLYrLYrSerGly	376
Q	y	1365	CTGAATAGAGAATTAACCAACACTGACCTGATCTCTGAGAAATATTACGCTTAAAAAA	1424
D	b	377	LeuAsnGInGInIleAsnAsnThrAspLeuIleSerGInLYrTYrValaLeuLYrLYr	396
Q	y	1425	GGGGAAAAAGCCGATAGATCCCTTTGATCGGAGTCACTTGAACTGTTCACCATCAAAATAC	1484
D	b	397	GlyGInLYrProTYrAspProPheAspArgSerhIleuLYrLYrPheThrIleLYrTYr	416
Q	y	1485	GTTGATGTGCATACCAACGAATTGCTAAAAAGTAGAGAGCTCTTAAACAGCTTAGCGAAGCT	1544
D	b	417	ValAspValAspThrAsnGInLeuLeuLYrSerGInGInLeuLeuThrAlaSerGInLYr	436
Q	y	1545	AACCTTACCTTCAGAGATTTATAGCATCCCGGATATAGGCTTAACTACTCTTCAACAAT	1604
D	b	437	AsnLeuAspPheAspArgAspLeuTYrAspProArgAspLYrAlaLYrLeuLeuTYrAsnAsn	456
Q	y	1605	CTCGATGCTTTTGGTATTATGACTATPACTTAACTGGAAGATGAGATTAATCAAGAT	1664
D	b	457	LeuAspAlaPheGlyIleMetAspTYrThrLeuThrGlyLYrValaGInAspAsnhIAsp	476
Q	y	1665	GACACCAACCGTATCATACCGTTTATATGGGCAAGGACCCGAAAGAGAGAAATGCTAGC	1724
D	b	477	AspThrAsnArgIleIleThrValTYrMetGlyLYrAspArgProGInGInAlaSnAlaSer	496
Q	y	1725	TACCATTTAGCT 1736	
D	b	497	TYrThIleuAla 500	
RESULT 11				
ID	AAW94665 standard; protein; 414 AA.			
XX	AAW94665;			
AC				
DT	04-MAY-1999 (first entry)			
XX				
DE	Streptococcus equisimilis mutant streptokinase K59E.			
XX				
DE	Streptococcus equisimilis H46A; streptokinase; mutant; fibrinolytic;			

KM plasma clot; hydrolysis; haemolytic Streptococcus; plasminogen; plasmin;
KM serine protease; fibrin; blood clot; thrombolytic;
KM vascular thrombolytic symptom; acute myocardial infarction;
KM fibrinolysis; resistance.
OS Streptococcus dysgalactiae subsp. equisimilis.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 14 /note= "encoded by ACC"
FT
XX US5876999-A.
XX
XX 02-MAR-1999.
XX
XX 06-DEC-1995; 95US-00568393.
XX
XX 06-DEC-1995; 95US-00568393.
XX
XX (NASC-) NAT SCI COUNCIL.
XX
XX PI
XX Wu H;
XX
XX WPI; 1999-189643/16.
XX N-PSDB; MAX16633.
XX
XX Mutant streptokinase polypeptide - useful as plasmin-resistant
XX thrombolytic agent.
XX
XX Claim 4; Col 11-14; 17pp; English.
XX
XX The present invention describes a mutant streptokinase (SK) polypeptide
XX in which at least one of the amino acids in the Pro58-Lys59-Ser60-Lys61
XX segment of the corresponding native SK is replaced by another amino acid.
XX The present sequence is mutant SK K59E. SK is a secretory protein of
XX haemolytic Streptococcus able to activate human plasminogen (Hpi) to
XX plasmin (Hpi), which is a serine protease able to catalyse the
XX hydrolysis of fibrin in blood clots. The SK is useful as a thrombolytic
XX agent in the treatment of vascular thrombolytic symptoms such as
XX acute myocardial infarction. Compared with wild-type SK, the K59E mutant
XX is more resistant to degradation by human plasmin and is more effective
XX both in acting as a fibrinolytic agent and in activating human plasminogen
XX
XX
SQ Sequence 414 AA;

Alignment Scores:
Pred. No.: 3.99e-171 Length: 414
Score: 1980.00 Matches: 382
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.74% Mismatches: 0
Query Match: 53.14% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-12 (1-2096) x AAM94665 (1-414)
QY 588 ATTGCTGACCTGAGTGGCTGCTAGACCGTTCATGTCACACAGCCCAATTGTTGTT 647
Db 1 IleaIaGIyProGluTrpLeuLeuAspArgProSerValaAsnAsnSerGlnLeuValVal 20
QY 648 AGCGTTGCTGTAACCTGTTGAGGGGACGAATCAAGACATTAGTCTTAATTTTGAATC 707
Db 21 SerValaIaGIyThValGIuGIyThraSngIaAspIleSerLeuLysPheGluIle 40
QY 708 GATCTAACATCAAGCACTGCTCATGAGAGAAAGACAGAGGCTTAAGTCAAAATCA 767
Db 41 AspLeuThrSerArgProIaHisGIyLysTrpGluGlnGlyLeuSerProGluSer 60
QY 768 AAACCAATTGCTACTGATAGTGGCGCGATGTCAATAAATTGAGAAAGCTGACTTA 827
Db 61 LysProPheAlaThrAspSerGIyAlaMetSerHisLysLeuGluLysAlaAspLeu 80
QY 828 AAGCTATTCAAGAAACAATTGATCGCTAACGTCACAGTAACAGACACTTGAAGTC 887

Db 81 LysAlaIleGIuGlnGlnLeuIleAlaLeuValHisSerAsnAspArgTrpPheGluVal 100
QY 888 ATTGATTTTGCAGACGATGACACCATTAAGTATGATGAAACGGCAAGTCTTACTGTCAC 947
Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyValTyrPheAlaAsp 120
QY 948 AAAGATGCTGGTAACTCTTCCGACCCCAACCTGTCACAAATTTTGCATAGCGGACAT 1007
Db 121 LysAspGlySerValThrLeuProTrpGlnProValGlnIuPheLeuSerGlyHis 140
QY 1008 GTGCGGCTTAAAGACATATAAAGAAAAACAATCAAAACCAAGCAAGAAATGTTGATGTG 1067
Db 141 ValArgValArgProTrpLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
QY 1068 GAATATAGCTGTACAGTTTACCTCCCTTAAACCTGATGACGATTTCAAGCAAGTCTCAAA 1127
Db 161 GluTrpThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
QY 1128 GATACCTAAGCTATTGAAAACACATGCTATTCGGTGACACCATCATCTCAAGAAATTA 1187
Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGIyAspThrIleThrSerGlnGluLeu 200
QY 1188 GCTCAAGCAAAAGCAATTTTAAACAAAACCAACCAAGGCTATAGATTTATGAAACGTGAC 1247
Db 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisAspGlyTrpThrIleTrpGluArgAsp 220
QY 1248 TCCTCAATCGTCACTCATGACATATGTCATTTCCGTACGATTTTCAACATGATCAAGAC 1307
Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
QY 1308 TTTACTTACCGGTGTAAATAATCGGAAACAAAGCTTATAGATCAATATAAATCGTCTG 1367
Db 241 PheThrLysArgValLysAsnArgGluGlnAlaTrpArgIleAsnLysLysSerGlyLeu 260
QY 1368 AATGAAGAATAAACAACACTGACCTGATCTCGAGAAATATTACGCTTAATAAAGGG 1427
Db 261 AsnGluGlnIleAsnAsnThrAspLeuIleSerGluLysTrpValLeuLysLysGly 280
QY 1428 GAAGACCGTATGATCCCTTGTATGTCAGTCACTTGAATCTGTCAACATCAATACGT 1487
Db 281 GluLysProTrpAspProPheAspArgSerHisLysLysLeuThrIleLysTrpVal 300
QY 1488 GATGTCGATACCAAGCAATTTGCTTAAAGTAGAGACGCTTTAACAGCTACGAAACGTAC 1547
Db 301 AspValAspThrAsnGluLeuLeuLysSerGluGlnLeuThrAlaSerGluArgAsn 320
QY 1548 TTTAGACTGAGATTTATACGATCCTCGATAGAGGCTTAACCTCAACAAATCTC 1607
Db 321 LeuAspPheArgAspLeuTrpAspProArgAspLysAlaLysLeuLeuTrpAsnAsnLeu 340
QY 1608 GATGCTTTGTTATTTATGACATTAACCTTAACGTGAAGAAAGTAAAGATTAATCAAGATAC 1667
Db 341 AspAlaPheGlyIleLeuMetAspTrpThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
QY 1728 CATTAGCT 1736
Db 381 HisLeuAla 383

RESULT 12
AAR12522
ID AAR12522 standard; protein; 483 AA.
XX
XX AAR12522;
AC
XX
XX 25-MAR-2003 (revised)
DT 17-SER-1991 (first entry)
XX
XX Factor Xa-cleavable streptokinase-IBGR-hirudin.

XX Fusion protein; blood clotting; coagulation; fibrinolysis;
 KW antithrombotic; thrombolysis.
 XX Synthetic.
 OS
 FH Key Location/Qualifiers
 FT Protein 1..414
 FT Peptide /label= streptokinase
 FT /label= linker
 FT /note= "factor Xa cleavage site"
 FT Protein 419..483
 FT /label= hirudin
 PN W09109125-A.
 XX 27-JUN-1991.
 PD
 PF 07-DEC-1989; 89GB-00027722.
 XX
 PR 07-DEC-1989; 89GB-00027722.
 PR 07-DEC-1990; 90MO-GB001911.
 PA (BRBI-) BRITISH BIO-TECHNOLOGY LTD.
 XX
 PI Dawson KM, Hunter MG, Czaplewsk LG;
 XX
 DR WPI, 1991-208151/28.
 DR N-PSDB; AAQ12490.
 PT Fusion protein cleavage by blood clotting enzyme - for prodn. of
 PT fractions having greater antithrombotic activity for therapy and
 PT prophylaxis.
 XX
 XX Disclosure; Page 96; 115pp; English.
 PS
 XX
 CC The protein is a recombinant product of a gene fusion construct. The
 CC sequence of the synthetic hirudin HV-1 genes was designed based on the
 CC published amino acid sequence (Dodd J., et al FEBS Letters 165 180
 CC (1984)). The sequence of streptokinase was obtcd. from PCR amplified
 CC chromosomal DNA from S. equisimilis ATCC 10009 or ATCC 9642. The primers
 CC used for the PCR were based on the pub- lished DNA sequence of S.
 CC equisimilis strain HA6A (Malke, H., Roe, B., and Rerretti, J.J., Gene 34
 CC 357-362 [1985]). The two sequences were used to construct an expression
 CC vector in which the streptokinase gene is linked to the hirudin gene via
 CC a linking sequence encoding a cleavage site for factor Xa. The factor Xa
 CC is present at the site of the target thrombus so the active agents are
 CC released specifically at the place where clot formation is occurring. See
 CC also AAR12887-R12889, AAR12891-R12894 and AAR12885. (Updated on 25-MAR-
 CC 2003 to correct PA field.)
 CC
 XX
 SQ Sequence 483 AA;
 Alignment Scores:
 Pred. No.: 6,49e-171 Length: 483
 Score: 1978.00 Matches: 401
 Percent Similarity: 82.28% Conservative: 17
 Best Local Similarity: 78.94% Mismatches: 52
 Query Match: 53.09% Indels: 38
 DB: 2 Gaps: 8
 US-09-940-235-12 (1-2096) x AAR12522 (1-483)
 QY 588 ATTCGACGCTGAGTGGCTGCTAGACCGTCATCTGTCAACAAAGCGAAATGTTGTT 647
 Db 1 IleaAGlyProGlntrpLeuLeuAspAsgProSerValAsnAsnSerGlnLeuVal 20
 QY 648 AGCGTGTGCTACTGTTTGAGGAGCAATCAAGACATTAGTCTTAATTTTGAATC 707
 Db 21 SerValAlaGlyThrValGlnGlnThrAsnGlnAspIleSerLeuysPhePheGlnIle 40
 QY 708 GATCTAACATCAAGCACTGCTCATGAGGAGAAAGACAGACAGAGCTTAAGTCCAAATCA 767

Db 41 AspLeuThrSerArgProAlaHisGlyLysThrGlnGlnGlyLeuSerProLysSer 60
 QY 768 AAACCATTTGCTACTGATAGTGGCGGATGTCACATTAACCTTGAGAAAGCTGACTTACTA 827
 Db 61 LysProPheAlaTrpAspSerGlyAlaMetProHisLysLeuGlnLysAlaAspLeu 80
 QY 828 AAGCGTATTCAAGAACATTTGATGCTTAACGTCACAGTAACGACGACTTGAAGTC 887
 Db 81 LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspAspLysPheGlnVal 100
 QY 888 ATTGATTTTGCAAGCGATGCAACCATTTACTGATCCGAAACGCGAAGTCTACTTGTGAC 947
 Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
 QY 948 AAAGATGTTGGTGAACCTTCCGACCCACACTGTCTCAAGAAATTTTCTTAAGCGGACAT 1007
 Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnGlnPheLeuSerGlyHis 140
 QY 1008 GTGGCGGTGACCATTAAGTAAAGAAACCATTAACAAACGCAAAATCTGTGATGTG 1067
 Db 141 ValArgValArgProTyrLysGlnLysProIleGlnAsnGlnAlaLysSerValAspVal 160
 QY 1068 GAATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTCAAGCAGGCTCAAA 1127
 Db 161 GlnTyrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuLys 180
 QY 1128 GATACCTAAGCTATTGAAAACTAGCTACGTGGTGAACACCATCACATCTCAAGAAATCTA 1187
 Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnLeuLeu 200
 QY 1188 GCTACAGCACAACAGATTTTAAACAAACCAACCGCTTACGATTTAGAAAGCTGAC 1247
 Db 201 AlaGlnAlaGlnSerIleLeuAsnLysThrAspProGlyTyrThrIleTyrGlnAsp 220
 QY 1248 TCCCAATCGTCATCTGATGACATGACATTTCCGTACGATTTTACCAATGATCAAG 1307
 Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGln 240
 QY 1308 TTACTTACCGTGTGTTAAAAATCGGAAACAGCTTATAGATCAATTAATAAAATCGTGTG 1367
 Db 241 PheThrLysHisValLysAsnArgGlnGlnAlaTyrGlnIleAsnLysLysSerGlyLeu 260
 QY 1368 AATGAAGAAATTAACAAACACTGACCTGATCTCTGAGAAATTTACGCTTTAAAAAGG 1427
 Db 261 AsnIleGlnIleLeuAsnThrAspLeuIleSerGlnLysTyrTyrValLeuLysLysGly 280
 QY 1428 GAAAGCGGTATGATCCCTTGAATCGGACGACCTGAAACGTTCACCATTAATACGT 1487
 Db 281 GlnLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
 QY 1488 GATGTCGATACCAAGCAATGCTTAATAAGTAGAGAGCTCTTAAGAGTGAAGCGTAAC 1547
 Db 301 AspValAsnThrAsnGlnLeuLeuLysSerGlnGlnLeuLeuThrAlaSerGlnAsn 320
 QY 1548 TTAGACTTCAGAGATTATACGATCTCGTATAGGCTAAACCTTCAACAATCTC 1607
 Db 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLysTyrAsnAsnLeu 340
 QY 1608 GATGCTTTTGGTATTATAGCACTATACCTTAACTGAAAAGTAGAGATATACAGATGAC 1667
 Db 341 AspAlaPheGlyIleMetAspTyrThrLeuThrIleLysValGlnAspAsnHisAspAsp 360
 QY 1668 ACCAAGCGTATCAATACGTTTATATAGGGAAGGACCCGAAGGAGGAATGCTAGCTAC 1727
 Db 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGlnGlnLysAsnAlaSerTyr 380
 QY 1728 CATTTAGCTGTGTGTGCGCAGCGCAAGATTGTCACCAATAGCTGAAGTCTTTGAT 1787
 Db 381 HisLeuAla-----TyrAspLysAspArgTyrThrGlnGlnGlnValTyrSer 398
 QY 1788 CATGCTGCGGAGCTTCTTATGTGCTCGAGAGAAACGTGGAGAAACCTTACCAAGGCTGG 1847
 Db 1788

Db 399 TyrLeuArgTyrThrGlyThrProIleProAspAsnProAsnAspIleGluGlyArg 418
 QY 1848 ATGATG---GTAGATTGACT-----TGCTGGAGAGAGGAGCGGA 1886
 Db 419 ValValTyrThrAspCysThrGlySerGlyGlnAsnLeuCysLeuGlySerAsn 438
 QY 1887 GCACATCACTTGACATCTTAGAAATAGATGCAACGATCAGACACAAAGACATCTTAGA 1946
 Db 439 -----ValCysGlyGlnGlnIleAsnIleCys-----Ile 447
 QY 1947 ATTGGAGACACCTGGAGAGAGAGATTAATCGAGAAACCTGCTCCAGTCATCTGCACA 2006
 Db 448 LeuGlySerAspGlyGlyIleAsn-----GlnCysVal---Thr 459
 QY 2007 GGCACGCGCCGAGAGAGAGTGAAGTGAAGAGCAGACCTGTGGAGACCAATGAGC 2066
 Db 460 GlyGlyGly-----ThyProIleProGlnSerHis 469
 QY 2067 GGATCTGGCCCTTCACCGATGT 2090
 Db 470 AsnAspGlyAspPheGluGluIle 477
 RESULT 13
 ID AAW21726 standard; protein; 1194 AA.
 AC AAW21726;
 DT 17-OCT-2003 (revised)
 DT 01-OCT-1997 (first entry)
 DE Streptokinase/maltose binding protein fusion protein, RSK.
 KM Plasminogen-binding fragment; streptokinase; degradation; MBP;
 KM Chromolytic agent; blood clot; bolus; maltose-binding protein.
 OS Streptococcus dysgalactiae subsp. equisimilis.
 FH Key location/Qualifiers
 FT Protein 1..381
 FT /label= Maltose binding protein
 FT /note= "acts as blocking group"
 FT Protein 382..1194
 FT /label= Streptokinase
 FN W09641883-A1.
 PD 27-DEC-1996.
 PE 07-JUN-1996; 96WO-US009640.
 PR 09-JUN-1995; 95US-00488940.
 PA (HARD) HARVARD COLLEGE.
 PI Reed GL;
 WP1; 1997-065469/06.
 Modified forms of streptokinase resistant to enzymatic cleavage - useful
 as thrombolytic agents in treating thrombosis and in medical equipment.
 Example 1; Page 11-12; 65pp; English.
 This sequence represents a fusion protein between maltose-binding protein
 and the plasminogen-binding fragment of streptokinase. This fusion
 protein was used in the design of modified streptokinase has an in vitro
 degradation rate at least 2 times slower than that of native
 streptokinase. Compounds containing modified streptokinases are
 specifically used as thrombolytic agents for dissolving blood clots in
 vivo in a mammal, preferably at a dose of 20000 U/kg, optionally as a
 bolus rather than by continuous infusion. (Updated on 17-OCT-2003 to
 standardise OS field)

XX SQ Sequence 1194 AA;
 Alignment Scores:
 Pred. No.: 1,31e-170 Length: 1194
 Score: 1976.50 Matches: 389
 Percent Similarity: 97.50% Conservative: 1
 Best Local Similarity: 97.25% Mismatches: 5
 Query Match: 53.05% Indels: 5
 DB: 2 Gaps: 3
 US-09-940-235-12 (1-2096) x AAW21726 (1-1194)
 QY 546 CAGACCAATCGAGC-----GGATCTGGCCCTTCACCGATGTTGTAATGCTGGA 596
 Db 366 GlnThrAsnSerSerSerValProGlyArgGlySerIle---GluGlyArgIleAlaIle 384
 QY 597 CTTGAGTGGCTGTAGACCGTCCATCTGTCAACAACGCAATTGTTGTTAGCGTTGCT 656
 Db 385 ProGluThrLeuLeuAspArgProSerValAsnAsnSerGlnLeuValSerValAla 404
 QY 657 GGTACTGTGAGGGAGACGAATCAAGACATTAAGTTAATTTTGAATGATCTACA 716
 Db 405 GlyThrValGluGlyThrAsnGlnAspIleSerLeuIlePheGluIleAspLeuThr 424
 QY 717 TCAAGACCTGCTGATGAG 776
 Db 425 SerArgProAlaHisGlyGlyLysTrpGlnGlnIleLeuSerProIleSerLeuPhe 444
 QY 777 GCTACTGATGATGGCGGATGTCATCAATTAAGAGAGAGAGAGAGAGAGAGAGAGAG 836
 Db 445 AlaThrAspSerGlyAlaMetSerHisIleLeuGlnIleValAlaSerLeuLeuValAlaIle 464
 QY 837 CAAGACAAATGATTCGTTAAGTTCACAGTCAAGACAGACATCTTGAGGTCATTTT 896
 Db 465 GlnGlnIleLeuIleAlaAsnValHisSerAsnAspAspIlePheGluValIleAspPhe 484
 QY 897 GCAAGCATGACCAACCATTAATCTGATTCGAAGAGAGAGAGAGAGAGAGAGAGAGAG 956
 Db 485 AlaSerAspAlaThrIleTrpAspArgAsnGlyLysValIlePheAlaAspLysAspGly 504
 QY 957 TCGGTAACTTGGCCGAGCCCAACCTGTCCAGAAATTTTGGTAAAGCGGACATGTGCGGTT 1016
 Db 505 SerValThrLeuProThrGlnProValGlnGlnIlePheLeuLeuSerGlyHisValArgVal 524
 QY 1017 AGACATATAAAGAAAAACCAATACAAACCAAGCAAGCAATCTGTTGATGGAAATTA 1076
 Db 525 Arg---TyrIleGlnIleProIleGlnAsnGlnAlaIleAspValAlaGluTyrThr 543
 QY 1077 GTACAGTTTACTCCCTTAAACCTGTAGAGATTTTGACACCAAGTCTCAAGATTAAC 1136
 Db 544 ValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuLysAspThrLys 563
 QY 1137 CTATTGAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTAAGTCAAGACA 1196
 Db 564 LeuLeuLysThrLeuAlaIleGlyAspThrIleHisSerGlnGlnLeuLeuAlaGlnAla 583
 QY 1197 CAAGCATTTTAAACAAAAACCAACCAAGGCTATACGATTTATGAACGTGACTCAATC 1256
 Db 584 GlnSerIleLeuLeuLysAsnHisProGlyTyrThrIleTyrGlnIleAspSerSerIle 603
 QY 1257 GTCACTCATGACATGACATTTTCCGTACGATTTTACCATGATGCAAGATTTACTTAC 1316
 Db 604 ValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlnPheThrTyr 623
 QY 1317 CGTGTAAATAATCGGAGAAAGGCTTATAGATCAATATAAATCGTGTGAATGAAGAA 1376
 Db 624 ArgValLysAsnArgGlnGlnAlaIleTyrArgIleAsnLysSerGlyLeuAsnGlnGln 643
 QY 1377 ATAAACAACATGACCTGATCTCTGAGAAATATATTCCTTAAAAAAAGGGAAGCCG 1436
 Db 644 IleAsnAsnThrAspLeuIleSerGlnLysTyrTyrValLeuLysGlyGlnLysPro 663

QY	1437	TATGATCCCTTTGATCGAGTCACTTGAACGTTCACACATCAATACGTTGATGCAT	1438
QY	1437	TATGATCCCTTTGATCGAGTCACTTGAACGTTCACACATCAATACGTTGATGCAT	1439
Db	664	TyrAspProPheAspArgSerHisIeuIysIeuPheTrpIleIysTyrValAspValAsp	683
QY	1497	ACCAACGAAATGCTTAAAAAGTAGCAGCGCTTTAACAGCTAGCGAAGCTTACTGAGCTTC	1556
Db	684	ThrAsnGluIleuIleuIysSerGluIleuIleuThrIleAsnGluArgAsnIleuAspHe	703
QY	1557	AGAGCTTTATACGATCCCTCGTATAGGCTAAAGCTTACTCTTACACAAATCTCGATGCTTTT	1616
Db	704	ArgAspLeuTyrAspProArgAspIysIleIuIysIleuTyrAsnAsnIleuAspAlaPhe	723
QY	1617	GCTATTATGAGCACTTATACCTTAACTCGAAAAGTAGAGGATATTCACGATGACCAACCGCT	1676
Db	724	GlyIleMetAspTyrThrIleuThrIlyIysValGluAspAsnHisAspAspThrAsnArg	743
QY	1677	ATCATTAACCGTTTATATATGGCAGCAGCCGAAAGAGAGAAATGCTAGCTTACATTAGCT	1736
Db	744	IleIleThrValTyrMetGlyIysArgProGluGluAsnAlaSerTyrHisIleuAla	763
RESULT 14			
ID	AAK12891	AAK12891 standard; protein; 499 AA.	
AC	AAK12891;		
DT	24-OCT-2003	(revised)	
DT	25-MAR-2003	(revised)	
DT	17-SEP-1991	(first entry)	
XX			
DE	Streptokinase fused to yeast alpha factor secretion sequence.		
XX			
KM	Fusion protein; blood clotting; coagulation; fibrinolysis;		
KM	antithrombotic; thrombolysis; streptokinase.		
XX			
OS	Streptococcus dysgalactiae subsp. equisimilis; ATCC 9542 or ATCC 100009.		
XX			
FX			
FT	Key	Location/Qualifiers	
FT	Peptide	1..85	
FT		/label= pre-pro alpha factor secretion sequence	
FT		/note= "S. cerevisiae"	
FT	Protein	86..499	
FT		/label= mature streptokinase	
XX			
FN	W09109125-A.		
PD	27-JUN-1991.		
XX			
PF	07-DEC-1989;	89GB-00027722.	
XX			
XX	07-DEC-1989;	89GB-00027722.	
FR	07-DEC-1990;	90WO-GB001911.	
XX			
PA	(BRBL-) BRITISH BIO-TECHNOLOGY LTD.		
XX			
PI	Dawson KM, Hunter MG, Czapliewsk LG;		
XX			
DR	WPI, 1991-208151/28.		
DR	N-PSDB; AAQ12158.		
XX			
PT	Fusion protein cleavage by blood clotting enzyme - for prodn. of		
PT	fibrinons having greater antithrombotic activity for therapy and		
PT	prophylaxis.		
XX			
PS	Disclosure; Page 86; 115pp; English.		
XX			
CC	The streptokinase sequence was obtd. from PCR amplified chromosomal DNA		
CC	from S. equisimilis (Lancetfield's Gp C) ATCC 10009 or ATCC 9642 (the		
CC	primers used for the PCR were based on the published DNA sequence of S.		
CC	equisimilis strain H46a (Malke, H., Roe, B., and Perretti, J.J., Gene 34		
CC	357-362 [1985]). The gene was fused to DNA encoding the yeast alpha		
CC	factor pre-pro-secretion sequence in an expression vector, pSMD1/152, for		
CC	prodn. of recombinant strepto- kinase in S. cerevisiae strain BU2168. See		

[illegible]

```

Db      |||
349  IleAsnAsnThrAspLeuIleSerGluYrYrValIleuYsGlyGluYsPro 368
Qy      1437 TATGATCCCTTGGATCGAGTCATGGAACCTTTCACCATCAAAATACGTGATCGCAT 1496
Db      369 TyrAspProPheAspArgSerHisLeuYsLeuPheThrIleYsYrValAspValAsn 388
Qy      1497 ACCAAGCAATTCGCTAAAGAGAGAGAGCTTCTTACAGCTAGCAAGCTTAAGTCACTTC 1556
Db      389 ThrAsnGluLeuLeuYsSerGluGlnLeuThrAlaSerGluThrAsnLeuAspPhe 408
Qy      1557 AGAGATTTATACGATCCTCGTATGAAGCTAAAGCTATCTTACAAACATCTCGATGCTTTT 1616
Db      409 ArgAspLeuYrAspProArgAspYsAlaYsLeuLeuYrAsnAsnLeuAspAlaPhe 428
Qy      1617 GGTATTATGACTATACCTTAACCTGAAAGATGAGAGATATACGATGACACCAACCGT 1676
Db      429 GlyIleMetAspYrThrIleuThrGlyYsValGluAspAsnHisAspAspThrAsnArg 448
Qy      1677 ATCATTAACGGTTTATATGAGGCAAGGAGCCGGAAGAGAGATGCTATGCTATTAAGCT 1736
Db      449 IleIleuValYrMetGlyYsArgProGluGlyGluAsnAlaSerYrHisLeuAla 468

RESULT 15
AAW21728
ID      AAW21728 standard; protein; 813 AA.
XX
AC      AAW21728;
XX
DT      17-OCT-2003 (revised)
DT      01-OCT-1997 (first entry)
XX
DE      Wild type plasminogen-binding fragment of Streptokinase.
XX
KW      Plasminogen-binding fragment; streptokinase; degradation; MBP;
KW      thrombolytic agent; blood clot; bolus; maltose-binding protein.
XX
OS      Streptococcus dysgalactiae subsp. equisimilis.
XX
XX      MO9641883-A1.
XX
PD      27-DEC-1996.
XX
PE      07-JUN-1996; 96WO-US009640.
XX
PR      09-JUN-1995; 95US-00488940.
XX
PA      (HARD ) HARVARD COLLEGE.
XX
PI      Reed GL;
XX
DR      WPI; 1997-065469/06.
XX
PT      Modified forms of streptokinase resistant to enzymatic cleavage - useful
PT      as thrombolytic agents in treating thrombosis and in medical equipment.
XX
PS      Example 1; Page 12-13; 65pp; English.
XX
XX      This sequence represents the wild type plasminogen-binding fragment of
XX      streptokinase. This fragment was used in the design of a modified
XX      streptokinase has an in vitro degradation rate at least 2 times slower
XX      than that of native streptokinase. Compounds containing modified
XX      streptokinases are specifically used as thrombolytic agents for
XX      dissolving blood clots in vivo in a mammal, preferably at a dose of 20000
XX      U/kg, opt. as a bolus rather than by continuous infusion. (Updated on 17-
XX      OCT-2003 to standardise OS field)
XX
SQ      Sequence 813 AA;

```

Alignment Scores: 9.08e-170 Length: 813
 Pred. No.: 1966.50 Matches: 382
 Score: 99.74% Conservative: 0
 Percent Similarity: 99.74%

```

Best Local Similarity: 99.74% Mismatches: 0
Query Match: 52.78% Indels: 1
DB: 2 Gaps: 1
US-09-940-235-12 (1-2096) x AAW21728 (1-813)

Qy      588 ATTGCTGACCTGAGTGGCTGTAGACCCGTCATCTGTCAACAACAGCAATGGTGGT 647
Db      1 IleAlaGlyProGluThrPheLeuAspArgProSerValAlaAsnSerGlnLeuVal 20
Qy      648 AGCGTGTCTGTACTGTGTGAGGAGCAATCAAGACATTAGCTTAAATTTTGAATC 707
Db      21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuYsPhePheGluIle 40
Qy      708 GATCTAACCTCAGACTGCTGTATGAGGAAACAACAAGCAAGCTTAAGTCCAAATCA 767
Db      41 AspLeuThrSerArgProAlaHisGlyGlyYsThrGluGlnGlyLeuSerProYsSer 60
Qy      768 AAACATTTGCTACTGATAGTGGCGGATGCTCAATMACTTGAGAAAGCTTACTTA 827
Db      61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluYsAlaAspLeu 80
Qy      828 AAGGCTATTCAAGAACAAATTGATCGCTAACGTCACAGTAAACAAGCTTATGAGTGC 887
Db      81 LysAlaAlaGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspAspYrPheGluVal 100
Qy      888 ATTGATTTTGCAGAGCGATGACACCATTTACTGATCGAAACGGCAAGGTCTACTGAC 947
Db      101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyYsValYrPheAlaAsp 120
Qy      948 AAGATGGTTCCGTTAAGCTTGGCCGACCCCAACCGTCCCAAGATTTTGGTAAAGGACAT 1007
Db      121 LysAspGlySerValThrLeuProGlnProGlnProValGlnIlePheLeuSerGlyHis 140
Qy      1008 GTGCGGGTTAGACCATATATAAGAAAAACAATCAAAACAAGGAAATCTGTGATGTC 1067
Db      141 ValArgValArg---YrLysGluYsProIleGlnAsnGlnAlaYsSerValAspVal 159
Qy      1068 GAATATACCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTGACAGCAAGCTCTCAA 1127
Db      160 GluYrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuYs 179
Qy      1128 GATACTAAGCTATTTGAAAAACATAGCTATGCGAGACCATCATCATCAAGATTACTA 1187
Db      180 AspThrLysLeuLeuYsThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 199
Qy      1188 GCTCAAGCAACAAGCAATTTTAAACAACAAACAACCCAGGCTATACGATTTTGAACGTGAC 1247
Db      200 AlaGlnAlaGlnSerIleLeuAsnLeuYsAsnHisProGlyYrThrIleYrGluArgAsp 219
Qy      1248 TCCTCAATGCTCACTCATGACAAATGACATTTTCGTCATGATTTTACCATGATCAAGAG 1307
Db      220 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 239
Qy      1308 TTACTTACCCGTTTAAAAATCGGAAACAAGCTTATAGATCAATAAATAATCGGCTG 1367
Db      240 PheThrYrArgValYsAsnArgGlnGlnAlaYrArgIleAsnYsLysSerGlyLeu 259
Qy      1368 AATGAAGAATAATAACAACACTGACCTGATCTGTGAGAAATATTACGCTTAAAAAGGG 1427
Db      260 AsnGlnGlnIleAsnAsnThrAspLeuIleSerGluYsYrYrValIleuYsGlyGly 279
Qy      1428 GAAAGCCCGTATGATCCCTTATGCGACGCTCACTTGAACCTGTTACCATCAATACGTT 1487
Db      280 GluYsProYrAspProPheAspArgSerHisLeuYsLeuPheThrIleYsYrVal 299
Qy      1488 GATTCGATACCAACGAATTCCTAAAGAGAGAGCTTTAACAGCTTACGAAACGTAAC 1547
Db      300 AspValAspThrAsnGlnLeuLeuYsSerGlnGlnLeuThrAlaSerGluArgAsn 319
Qy      1548 TTAGACTTCAGAGATTTATACGATCCTCGGATAGAGCTTAACTACTCAACAACTCTC 1607
Db      320 LeuAspPheArgAspLeuYrAspProArgAspYsAlaYsLeuLeuYrAsnAsnLeu 339

```


OY	1608	GATGCTTTGGTATTATGACTATACCTTA	CTGAAAGTAGAGATAATCACGATGAC	1667
Db	340	AspAlaPheGlyIleMetAspTyrThrIleu	ThrGlyLysValGluAspAsnHisAspAsp	359
OY	1668	ACCAACCGTATCATACCGTTTATATGGG	CACGACCGAAGAGAGAAATGCTAGCTAC	1727
Db	360	ThrAsnArgIleIleThrValTyrMetGly	LysArgProGluGlyGluAsnAlaSerTyr	379
OY	1728	CATTAGCT	1736	
Db	380	HisLeuAla	382	

Search completed: January 28, 2006, 02:09:01
 Job time : 167.802 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 28, 2006, 01:49:55 ; Search time 29.9006 Seconds
(without alignments)
11590.966 Million cell updates/sec

Title: US-09-940-235-12

Perfect score: 3726
Sequence: 1 cgaagaccatcatgtctgtc.....ccttcacgcatgtcgttag 2096

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 572060 seqs, 82675679 residues
Total number of hits satisfying chosen parameters: 1144120

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
MODEL=frame+ n2p.model -DEV=xlD
Q=/cgn2_1/USPTO_spool_p/US09940235/runat_27012006_144219_27602/app_query.fasta_1.7708
DB=Issued_Patents_AA -QFMT=fastan -SUFFIX=rai -BINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsun62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09940235 -ECGN 1.1.159 -runat_27012006_144219_27602 -NCPU=6 -ICPU=3
-NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_AA:*
1: /cgn2_6/prodata/1/iaa/5 COMB.pdp:*
2: /cgn2_6/prodata/1/iaa/6 COMB.pdp:*
3: /cgn2_6/prodata/1/iaa/H COMB.pdp:*
4: /cgn2_6/prodata/1/iaa/PCTUS COMB.pdp:*
5: /cgn2_6/prodata/1/iaa/RE COMB.pdp:*
6: /cgn2_6/prodata/1/iaa/backfile1.pdp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1994	53.5	795	2	US-09-211-542A-2
2	1991	53.4	483	1	US-07-854-596B-43
3	1984	53.2	413	2	US-10-360-101-264
4	1984	53.2	414	2	US-09-211-542A-6
5	1984	53.2	414	6	5240845-1
6	1984	53.2	440	1	US-08-560-098A-52
7	1978	53.1	483	1	US-07-854-596B-47
8	1976.5	53.0	1194	1	US-08-488-940-1
9	1969	52.8	499	1	US-07-854-596B-28
10	1966.5	52.8	813	1	US-08-488-940-3
11	1958	52.5	415	1	US-07-854-596B-26
12	1958	52.5	435	1	US-07-854-596B-19

13	1958	52.5	440	1	US-07-854-596B-15	Sequence 15, Appl
14	1958	52.5	859	1	US-07-854-596B-35	Sequence 35, Appl
15	1947.5	52.3	1194	1	US-08-488-940-18	Sequence 18, Appl
16	1946.5	52.2	1194	1	US-08-488-940-17	Sequence 17, Appl
17	1928	51.7	413	1	US-08-759-559-12	Sequence 12, Appl
18	1928	51.7	413	2	US-09-294-457-12	Sequence 12, Appl
19	1928	51.7	413	2	US-09-919-703-12	Sequence 12, Appl
20	1902	51.0	384	2	US-09-374-038-4	Sequence 4, Appl
21	1902	51.0	384	2	US-09-658-179-4	Sequence 4, Appl
22	1898	50.9	372	2	US-09-374-038-3	Sequence 3, Appl
23	1898	50.9	372	2	US-09-658-179-3	Sequence 3, Appl
24	1897	50.9	800	1	US-08-488-940-4	Sequence 4, Appl
25	1897	50.9	1181	1	US-08-488-940-2	Sequence 2, Appl
26	1890.5	50.7	747	1	US-07-854-596B-40	Sequence 40, Appl
27	1887	50.6	401	2	US-09-374-038-1	Sequence 1, Appl
28	1887	50.6	401	2	US-09-658-179-1	Sequence 1, Appl
29	1887	50.6	413	2	US-09-374-038-2	Sequence 2, Appl
30	1887	50.6	413	2	US-09-658-179-2	Sequence 2, Appl
31	1875	50.3	369	1	US-07-854-596B-31	Sequence 31, Appl
32	1698.5	45.6	736	2	US-09-211-542A-4	Sequence 4, Appl
33	1686	45.2	356	2	US-09-211-542A-12	Sequence 12, Appl
34	785	21.1	150	2	US-09-211-542A-14	Sequence 14, Appl
35	640.5	17.2	2231	1	US-08-153-799-16	Sequence 16, Appl
36	640.5	17.2	2386	1	US-10-360-101-235	Sequence 12, Appl
37	638.5	17.1	2355	2	US-10-360-101-235	Sequence 235, App
38	638.5	17.1	2386	2	US-09-961-403-1	Sequence 1, Appl
39	636	17.1	2446	1	US-08-551-356-2	Sequence 2, Appl
40	636	17.1	2446	4	PCT-US93-12687-2	Sequence 2, Appl
41	632	17.0	2324	1	US-08-283-857-1	Sequence 1, Appl
42	632	17.0	2324	4	PCT-US95-09819-1	Sequence 1, Appl
43	627	16.8	2327	6	5455158-1	Patent No. 5455158
44	283	7.6	286	2	US-09-078-091-4	Sequence 4, Appl
45	275	7.4	286	2	US-09-078-091-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-09-211-542A-2
; Sequence 2, Application US/09211542A
; Patent No. 6210667
; GENERAL INFORMATION:
; APPLICANT: Reed, Guy L.
; TITLE OF INVENTION: BACTERIAL FIBRIN-DEPENDENT PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: BROMBERG & SUNSTEIN, LLP
; STREET: 125 Summer Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/211,542A
; FILING DATE: 15-December-1998
; CLASSIFICATION: 1653
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/069,497
; FILING DATE: 15-December-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Attorney, Strimpel, Harriet M.
; REGISTRATION NUMBER: 37,008
; REFERENCE/DOCKET NUMBER: 1874/111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)443-9292
; TELEFAX: (617)443-0004
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:

LENGTH: 795 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-211-542A-2

Alignment Scores:

Pred. No.:	1.65e-190	Length:	795
Score:	1994.00	Matches:	390
Percent Similarity:	97.75%	Conservative:	1
Best Local Similarity:	97.50%	Mismatches:	5
Query Match:	53.52%	Indels:	4
		Gaps:	2

US-09-940-235-12 (1-2096) x US-09-211-542A-2 (1-795)

```

QY 546 CAGACACATCGAGC-----GGATCTGGCCCTTCAGCCGATGTCGTATGCTGCA 596
Db 366 GlnThrAsnSerSerValProGlyArgGlySerIle---GluGlyArgIleAlaGly 384
QY 537 CCGAGTGCGCTCTGAGCCGCTCATCTGTCAACACAGCCAAATGCTTGTAGCGTTGCT 656
Db 385 ProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuValValSerValAla 404
QY 657 GGTACTGTGAGGGGAGCAATGACATGACTTAAATTTTGAATTCGATCTTAACA 716
Db 405 GlnTrpValGluGlyThrAsnGlnAspIleSerLeuLysPheGluValIleAspLeuThr 424
QY 717 TCACGACCTGCTCATGAGGAGAAAGACAGAGCAAGGCTTAAGTCCAAATCAAAACCATTT 776
Db 425 SerArgProIleAsnIleGlyGlySerThrGluGlnGlyLeuSerProLysSerLysProPhe 444
QY 777 GCTACTGATAGTGGCGCGCATGTCACATAACTTGAGAAAGCTGACTTAAAGCTTAT 836
Db 445 AlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeuLysAlaIle 464
QY 837 CAAGAACAAATTGATGCTTAACGTCCACAGTAACGACTACTTTGAGGTCATGATTTT 896
Db 465 GlnGlnGlnLeuIleAlaAsnValHisSerAsnAspAspLysPheGluValIleAspPhe 484
QY 897 GCAAGCGATGCAACCATTAAGTCAAAACGCAAGGCTTACTTTCTGACAAAGATGAT 956
Db 485 AlaSerAspAlaThrIleThrAspArgAsnGlyLysValLysPheAlaAspLysAspGly 504
QY 957 TCGGTAACCTTGGCGACCCCAACCTGTCCAAAGATTTTCTTAAAGCGGACATGCGCGGT 1016
Db 505 SerValThrLeuProThrGlnProValGlnGlnPheLeuLeuSerGlyHisValArgVal 524
QY 1017 AGACCATATTAAGAAACCATTAACAAACCAAGCAAGTCTGTTGATGTGGAATTAAT 1076
Db 525 ArgProLysLysGlnLysProLleGlnAsnGlnAlaLysSerValAspValGluLysTrpThr 544
QY 1077 GTACAGTTTACTCCCTTAAACCTGATGACATTTTCAGACCAAGGCTTCAAGATACATAG 1136
Db 545 ValGlnPheThrProLeuAsnProAspAspAspPheAspProGlyLeuLysAspTrpLys 564
QY 1137 CTATTGAAACCACTAGTATGCGTGAACACATCAATCTCAAGAAATTAAGTCAAGCA 1196
Db 565 LeuLeuLysThrLeuAlaIleGlyAspTrpIleThrSerGlnGluLeuLeuAlaGlnAla 584
QY 1197 CAAAGCATTTTAAACAAACCAACCAAGGCTTACGATTTTAAAGAGTGAATCTCAATC 1256
Db 585 GlnSerIleLeuAsnLysAsnHisAspProGlyLysTrpIleLysGlnLysArgAspSerIle 604
QY 1257 GTCACTCATGACAAATGACATTTTCGTAAGATTTTACAAATGATCAAGGTTTACTTAC 1316
Db 605 ValThrHisAspAsnAspIlePheAspTrpIleLeuProMetAspGlnGluPheTrpLys 624
QY 1317 CGTGTAAATAATCGGAAACAAAGCTTAAGATCAATAAATAATCGTGCTGAATGAAGA 1376
Db 625 ArgValLysAsnArgGluGlnAlaLysTrpArgIleAsnLysLysSerGlyLeuAsnGlnGlu 644
QY 1377 ATAAACAACACTGACCTGATCTCTGAGAAATATTTACGTCCTTAAATAAGGGAAAGCCG 1436

```

```

Db 645 IleAsnAsnThrAspLeuIleSerGlnLysTrpValLeuLysGlyGlnLysPro 664
QY 1437 TATGATCCCTTGTAGTCGACGTCACCTGAAACCTGTCCACATCAATACTGATGTCAT 1496
Db 665 TyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTrpValAspValAsp 684
QY 1497 ACCAACAATTTGCTTAAATAAGTACGACCTCTTAAACACTAGCCGAACCTTAAGCTTTC 1556
Db 685 ThrAsnLysLeuLeuLysSerGlnGlnLeuLeuThrAlaSerGlnLysAsnLeuAspPhe 704
QY 1557 AGAATTTATACGATCTCGTGAATPAAGCTTAACTTACTTACAAATCTCGATGCTTTT 1616
Db 705 ArgAspLeuLysTrpAspProAspAspLysAlaLysLeuLeuLysAsnAsnAspAlaPhe 724
QY 1617 GGTAATTATGACATATACCTTAACTGAAAGTAGAGATTAATCAAGTACCAACCGCT 1676
Db 725 GlyLeuMetAspTrpTrpThrLeuThrGlnLysValGluAspAsnHisAspAspThrAsnArg 744
QY 1677 ATCATTAACGCTTATATATGGGCAAGCCGACCGAAGAGAGAAATGCTTACCATTTAGCT 1736
Db 745 IleIleThrValLysMetGlyLysArgProGlnGlyGluAsnAlaSerTrpHisLeuAla 764

```

RESULT 2

US-07-854-596B-43

Sequence 43, Application US/07854596B
 Patent No. 5434073

GENERAL INFORMATION:

APPLICANT: Dawson, Keith M
 APPLICANT: Hunter, Michael G
 APPLICANT: Czaplowski, Lloyd G
 TITLE OF INVENTION: Proteins and nucleic acids
 NUMBER OF SEQUENCES: 73
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Dr. John J. McDonnell
 STREET: Ten South Wacker Drive, Suite 3000
 CITY: Chicago
 STATE: IL
 COUNTRY: USA

ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/854,596B
 FILING DATE: 03-JUN-1992

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
 NAME: McDonnell, John J
 REGISTRATION NUMBER: 26,949
 REFERENCE/DOCKET NUMBER: 92,337
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312-715-1000
 TELEFAX: 312-715-1234

TELEX: 910-221-5317

INFORMATION FOR SEQ ID NO: 43:

SEQUENCE CHARACTERISTICS:
 LENGTH: 483 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-07-854-596B-43

Alignment Scores:

Pred. No.:	2.5e-190	Length:	483
Score:	1991.00 <td>Matches:</td> <td>403</td>	Matches:	403
Percent Similarity:	84.29%	Conservative:	10
Best Local Similarity:	82.24%	Mismatches:	27
Query Match:	53.44%	Indels:	50
DB:	1	Gaps:	9


```

QY 768 AAACATTTGCTACTGATGATGCGGATGTCACATTAATCTTGAAAGGAGTGAATCTTACTA 827
Db 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLeuGluValAspLeu 80
QY 828 AAGGCTATTCAAGAAACAATTGATCGGTAAAGTCCACAGTAAACGACATCTTGAAGTC 887
Db 81 LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspGlyPheGluVal 100
QY 888 ATTGATTTTGGCAGGATGCAACCATTTACTGATCGAAACGGCAGGTCTTACTTGGTAC 947
Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyValTyrPheAlaAsp 120
QY 948 AAAGATGTTGGGTAACTTGGCCGACCCGACCGTCCAGAAATTTTGGTAAAGCGGACAT 1007
Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuSerGlyHis 140
QY 1008 GTGCGCGTTAAGCATATTAAGAAAAACCATTAACAAACCAAGCGAAATCTGTGATGTG 1067
Db 141 ValArgValArgProTyrGlySerGlyValSerProIleGlnAsnGlnAlaLysSerValAspVal 160
QY 1068 GAATTAATCTGACAGTTTACTCCCTTAAACCTGATGACGATTTTCAGACCGAGTCTCAAA 1127
Db 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuLys 180
QY 1128 GATCTAAGCTATTGAAAAACAATAGCTATCGGTGACACATCAATCAATCAATCAATCAATCA 1187
Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 200
QY 1188 GCTCAAGCAACAAGCATTTTAAACAACCAACCCGAGCTATACGATTTATGAACGTGAC 1247
Db 201 AlaGlnAlaGlnSerIleLeuLeuLysAsnHisProGlyTyrThrIleTyrGluAsp 220
QY 1248 TCCTCAATGCTACTCATGACATGACATATTTCCGTACGATTTTACCAATGATCAAGAG 1307
Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
QY 1308 TTTACTTACCGGTGTTAAAAATGCGGACAAAGCTTATAGATCAATAAAAATCTGCTCG 1367
Db 241 PheThrTyrArgValLysAsnArgGlnGlnAlaTyrArgIleAsnLysLysSerGlyLeu 260
QY 1368 AATGAAGAAATTAACAACAACATGACCTGATCTCTGAGAATATTTACGCTTAAAAAGG 1427
Db 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysGly 280
QY 1428 GAAAGCCGATGATCCCTTTGATCGCACTCACTTGAACCTGTCCACATCAATACGTT 1487
Db 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
QY 1488 GATGTCGATACCAAGCAATGCTTAAAAAGTGAACGCTTTTAAACGCTTACGCAACGTAAC 1547
Db 301 AspValAspThrAsnGluLeuLeuLysSerGlnGlnLeuThrAlaSerGluAsn 320
QY 1548 TTAAGCTCAGAGATTTATAGATCCTGCTGATTAAGGCTTAACTACTTCAACAATCTC 1607
Db 321 LeuAspPheAspLeuLysAspProAspAspLysAlaLysLeuLeuLysAsnLeu 340
QY 1608 GATGCTTTGGTATTATGACTATACCTTAACTGAAAAGTAGAGGATTAATCAAGTAGAC 1667
Db 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyValGluAspAsnHisAspAsp 360
QY 1668 ACCAACGCTATCATACCGTTTATATGGCAAGCGAACCGGAAGAGAGAAATGCTACTTAC 1727
Db 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluLysAsnAlaSerTyr 380
QY 1728 CATTAGCT 1736
Db 381 HisLeuAla 383

```

RESULT 4
 US-09-211-542A-6
 ; Sequence 6, Application US/09211542A
 ; Patent No. 6210667

```

; GENERAL INFORMATION:
; APPLICANT: Reed, Guy L.
; TITLE OF INVENTION: BACTERIAL FIBRIN-DEPENDENT PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROMBERG & SUNSTEIN, LLP
; STREET: 125 Summer Street
; City: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/211,542A
; FILING DATE: 15-December-1998
; CLASSIFICATION: 1653
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/069,497
; FILING DATE: 15-December-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Attorney, Strimpel, Harriet M.
; REGISTRATION NUMBER: 37,008
; REFERENCE/DOCKET NUMBER: 1874/111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 443-9292
; TELEFAX: (617) 443-0004
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 414 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULAR TYPE: protein
US-09-211-542A-6
Alignment Scores:
Pred. No.: 1,156-189 Length: 414
Score: 1984.00 Matches: 383
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 53.25% Indels: 0
Gaps: 0
US-09-211-542A-6 (1-2096) x US-09-211-542A-6 (1-414)
QY 588 ATTGCTGACCTGAGCGGTGCTGACACCGTCCATCTGTCGAACAACGCAATGCTTGT 647
Db 1 IleAlaGlyProGluTyrLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
QY 648 AGCGTGTGCTACTGCTGTGAGGGAAGCAATCAAGACATTAAGTCTTAAATTTTGAATC 707
Db 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 40
QY 708 GATTAACATCAACGACCTGCTCATGAGAGAAAGACAGAGAGGCTTAAGTCCAAAAATCA 767
Db 41 AspLeuThrSerArgProAlaHisGlyGlyLysThrGlnGlnGlnLeuSerProLysSer 60
QY 768 AAACATTTGCTACTGATGATGCGGATGTCACATTAATCTTGAAAGGAGTGAATCTTACTA 827
Db 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLeuGluValAspLeu 80
QY 828 AAGGCTATTCAAGAAACAATTGATCGGTAAAGTCCACAGTAAACGACATCTTGAAGTC 887
Db 81 LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspGlyPheGluVal 100
QY 888 ATTGATTTTGGCAGGATGCAACCATTTACTGATCGAAACGGCAGGTCTTACTTGGTAC 947
Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyValTyrPheAlaAsp 120
QY 948 AAAGATGTTGGGTAACTTGGCCGACCCGACCGTCCAGAAATTTTGGTAAAGCGGACAT 1007

```

Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnIleuPheLeuSerGlyHis 140
1008 GTGCGCGTTAGACCATATTAAGAAAAACAATACAAGGAAATCGTTGATGTG 1067
Db 141 ValArgValArgProGlyArgGlyLeuProIleGlnAsnGlnAlaLysSerValAspVal 160
Qy 1068 GAATATAGCTGATACGTTTACCTCCCTTAACCCCTGATGACGATTTGACAGCGCTGAAA 1127
Db 161 GluIyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
Qy 1128 GATTACTTAACCTATTTGAAAAACCTAGCTATCGGTGACCACTCATCTTCAAGAAATTA 1187
Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnIleuLeu 200
Qy 1188 GCTCAAGCAAAAGCATTTTAAACAAAACCAACCGGCTATCGATTTTGAACCGTAC 1247
Db 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyIyrThrIleIyrGlnArgAsp 220
Qy 1248 TCCTCAATCGTCACTGATGACATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 1307
Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
Qy 1308 TTTACTTACCGGTGTTAAAAATCGGGAACAAGCTTATAGATCAATAAAAATCGGTCTG 1367
Db 241 PheThrIyrArgValLysAsnArgGlnGlnAlaIyrArgIleAsnLysLysSerGlyLeu 260
Qy 1368 AATGAAGAATAAACAACAACGATGACCTGATCTCGAAGAAATATACGCTTCAAAAAAGG 1427
Db 261 AsnGlnGlnIleAsnAsnThrAspLeuIleSerGlnLysIyrIyrValLeuLysGly 280
Qy 1428 GAAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTCACATCAATAGCT 1487
Db 281 GluIyrProIyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysIyrVal 300
Qy 1488 GATGTCGATACCAACGATTTGCTAAAGAGTGAAGACGCTTTAACAGCTAGCGAAGCTAC 1547
Db 301 AspValAspThrAsnGlnLeuLeuLysSerGlnGlnLeuThrAlaSerGlnLysArgAsn 320
Qy 1548 TTAAGACTTCAAGATTTATACGATCTCGTGAATAGGCTTAACCTTCAACAATCTC 1607
Db 321 LeuAspPheArgAspLeuLysAspProArgAspLysAlaLysLeuLeuLysIyrAsnAsnLeu 340
Qy 1608 GATGCTTTGATATTATGAGCTATATCCTTAACCTGGAAGAGTAACTACGATGAC 1667
Db 341 AspAlaPheGlyIleLeuAspIyrThrLeuThrGlyLysValGlnAspAsnHisAspAsp 360
Qy 1668 ACCAACCCTATCATTAACCGTTTATATGGCAAGCGAACCGGAAGAGATGCTAGCTAC 1727
Db 361 ThrAsnArgIleIleThrValIyrMetGlyLysArgProGlnGlyGlnAsnAlaSerIyr 380
Qy 1728 CATTTAGCT 1736
Db 381 HisLeuAla 383

RESULT 5
5240845-1
; Patent No. 5240845
; APPLICANT: FUJII, SETSURO; TAKADA, KAORUKO.; KATANO, TAMIKI;
; MAJIMA, EIJI; OGINO, KOICHI; ONO, KENJI; SAKATA, YASUYO; UENOYAMA,
; TSUTOMU
; TITLE OF INVENTION: MUTATED STREPTOKINASE PROTEINS
; NUMBER OF SEQUENCES: 65
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/549,049
; FILING DATE: 06-JUL-1990
; SEQ ID NO.: 1
; LENGTH: 414
5240845-1

Alignment Scores:
Pred. No.: 1,156-189 length: 414
Score: 1984.00 Matches: 383

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 53.25% Indels: 0
DB: 6 Gaps: 0

US-09-940-235-12 (1-2096) x 5240845-1 (1-414)

Qy 588 ATTGCTGAGCACTGAGTGGCTCTGACCGCTCATCTGTGCAACAAGCAATGTTGTTGT 647
Db 1 IleAlaGlyProGlnIyrProIyrPheLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
Qy 648 AGCGTTCCGTGCTACTGTTGAGGGGACGAATCAAGACATTAAGCTTAAATTTTGAATC 707
Db 21 SerValAlaGlyThrValGlnGlyIyrThrAsnGlnAspIleSerLeuLysPhePheGlnIle 40
Qy 708 GATTAACATACGACACCTGCTCATGAGAGAAAGACAGCAAGGCTTAAGTCAAAATCA 767
Db 41 AspLeuThrSerArgProAlaHisGlyGlyLysThrGlnGlnIyrLeuSerProLysSer 60
Qy 768 AAACCATTTGCTGATGATGAGGCGGATGACATTAACCTTGAAGAAAGCTTACTTA 827
Db 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlnLysAlaAspLeuLeu 80
Qy 828 AAGGCTATTCAAGAACATTTGATGCTTACGTCACAGTAAACGACACTTACTTGAAGCTC 887
Db 81 LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspAspIyrPheGlnVal 100
Qy 888 ATGATATTTGCAAGCGATGCAACCATTAAGTGAAGCGGAAAGCTTACTTGTCTGAC 947
Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValIyrPheAlaAsp 120
Qy 948 AAGAATGCTTCGTTAAACCTTCCGACCCGACCTGTCAGAAATTTTGTCAAGGCAAT 1007
Db 121 LysAspLysSerValThrLeuProThrGlnProValGlnIyrPheLeuSerGlyHis 140
Qy 1008 GTGCGCGTTAGACCATATTAAGAAAAACAATACAAGGAAATCGTTGATGTG 1067
Db 141 ValArgValArgProIyrLysGlnLysProIleGlnAsnGlnAlaLysSerValAspVal 160
Qy 1068 GAATATAGCTGATACGTTTACCTCCCTTAACCCCTGATGACGATTTTGAACCGTCAAG 1127
Db 161 GluIyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
Qy 1128 GATTACTTAACCTATTTGAAAAACCTAGCTATCGGTGACCACTCATCTTCAAGAAATTA 1187
Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnIleuLeu 200
Qy 1188 GCTCAAGCAAAAGCATTTTAAACAAAACCAACCGGCTATCGATTTTGAACCGTAC 1247
Db 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyIyrThrIleIyrGlnArgAsp 220
Qy 1248 TCCTCAATCGTCACTGATGACATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 1307
Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
Qy 1308 TTTACTTACCGGTGTTAAAAATCGGGAACAAGCTTATAGATCAATAAAAATCGGTCTG 1367
Db 241 PheThrIyrArgValLysAsnArgGlnGlnAlaIyrArgIleAsnLysLysSerGlyLeu 260
Qy 1368 AATGAAGAATAAACAACAACGATGACCTGATCTCGAAGAAATATACGCTTCAAAAAAGG 1427
Db 261 AsnGlnGlnIleAsnAsnThrAspLeuIleSerGlnLysIyrIyrValLeuLysGly 280
Qy 1428 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCAACATCAATAGCT 1487
Db 281 GluLysProIyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysIyrVal 300
Qy 1488 GATGTCGATACCAACGATTTGCTAAAGAGTGAAGACGCTTTAACAGCTAGCGAAGCTAC 1547
Db 301 AspValAspThrAsnGlnLeuLeuLysSerGlnGlnLeuThrAlaSerGlnLysArgAsn 320
Qy 1548 TTAAGACTTCAAGATTTATACGATCTCGTGAATAGGCTTAACCTTCAACAATCTC 1607

Db 321 LeuAspPheArgAspLeuTyraSPProArGAspLysAlaLysLeuLeuTyraAsnLeu 340
Qy 1608 GATGCTTTGGTATTATGAGCTTACTTAACTGGAAAGTAAAGATATTCAGATGAC 1667
Db 341 AspAlaPheGlyLysLeuAspTyrrThrLeuThrglyLysValGluAspAsnHisAspAsp 360
Qy 1668 ACCAACCGTATCATTAACGTTTATATGAGAGCAAGCCGAGAGAGAGAAATGCTAGCTAC 1727
Db 361 ThrAsnArgIleIleThrValTyrrMetGlyLysArgProGluGlyGluAsnAlaSerTyrr 380
Qy 1728 CATTTAGCT 1736
Db 381 HisLeuAla 383
RESULT 6
US-08-560-098A-52
Sequence 52, Application US/08560098A
Patent No. 5976841
GENERAL INFORMATION:
APPLICANT: WNIENDT, Stephan
APPLICANT: HEINZEL-WIELAND, Regina
APPLICANT: STEFFENS, Gerd Josef
TITLE OF INVENTION: Proteins having fibrinolytic and
TITLE OF INVENTION: Coagulation-inhibiting Properties
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/560,098A
FILING DATE: 17-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P 44 40 892.7
FILING DATE: 17-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/42448
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 52:
SEQUENCE CHARACTERISTICS:
LENGTH: 440 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-560-098A-52
Alignment Scores:
Pred. No.: 1,19e-189 Length: 440
Score: 1984.00 Matches: 383
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 53.25% Indels: 0
Gaps: 0
US-09-940-235-12 (1-2096) x US-08-560-098A-52 (1-440)
Qy 588 ATTCGTGACCTGATGCTGCTAGACCGTCACTGTGCAACACAGCAATGCTGTT 647
Db 27 IleAlaGlyProGluTrpLeuLeuAspAlaGProSerValAsnAsnSerGlnLeuValVal 46

Qy 648 AGCGTTCGTGTACTGTGAGGGAGCAATCAAGACATTAAGTCTTAATTTTGAATATC 707
Db 47 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 66
Qy 708 GATCTAACATCAAGCACTGCTCATGAGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA 767
Db 67 AspLeuThrSerArgProAlaHisGlyGlySerThrGluGlnGlyLeuSerProLysSer 86
Qy 768 AAACCATTTTCTACTGATAGTGGCGCATGTCACTAACTTGAGAAAGCTGACTTACTA 827
Db 87 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 106
Qy 828 AAGGCTATTCAAGAACTTGAATGCTTAAGTCCAGATCAAGCAAGCTTCTTGAAGTTC 887
Db 107 LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspTyrrPheGluVal 126
Qy 888 ATGATTTTTCAGAGCAAGTCAACCATTTACATGCAAGCAAGGCAAGTCTTCTTGTGAC 947
Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlySerValTyrrPheAlaAsp 146
Qy 948 AAGATGCTTCGGTAACTTGGCCGACCAACCTGTCCAGAAATTTTGTCAAGCGGACAT 1007
Db 147 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 166
Qy 1008 GTGGCGTTAGACCATTAAGAAACCAATACCAACCAAGGAAATCTGTGATGTG 1067
Db 167 ValArgValArgProTyrrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 186
Qy 1068 GAATATATCTGATGATTTTACCTCTTAAACCTGATGACGATTTTCAGACGAGTCTCAAA 1127
Db 187 GluTyrrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuLys 206
Qy 1128 GATACTAAGCTATTGAAACACATAGCTATGGTGACCAATCATCTCAAGAAATTACTA 1187
Db 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 226
Qy 1188 GCTCAAGCAACAAGCATTTTAAACAAACCAACCCAGGCTTATACGATTTATGAACGTGAC 1247
Db 227 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisAspProGlyTyrrThrIleTyrrGluArgAsp 246
Qy 1248 TCCGCAATGCTCACTCAATGACAAATGACATTTTCCGTGCAATTTTCAACGATCAAGAG 1307
Db 247 SerSerIleValIleThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
Qy 1308 TTACTTACCGTGTAAATAATCGGAAACAAGCTTATAGATCAATAAATAATCGTGTG 1367
Db 267 PheThrTyrrArgValLysAsnArgGluGlnAlaTyrrArgIleLeuLysLysSerGlyLeu 286
Qy 1368 AATGAAGAATAAACAACATGACCTGATCTTGAGAAATTTACGTCCTTAAAAAGGG 1427
Db 287 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrrTyrrValLeuLysLysGly 306
Qy 1428 GAAAGGCGTATGATCCCTTGTATGCGAGTCACTTGAACGCTTCAACATCAATTAAGTT 1487
Db 307 GluLysProTyrrAspProPheAspArgSerHisLysLysLeuPheThrIleLysTyrrVal 326
Qy 1488 GATGTCGATCAACAAGATTTGCTAAAGAGAGAGCTTTAACGCTAGCGAAGCTTAAC 1547
Db 327 AspValAspThrAsnGlnLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
Qy 1548 TTAGACTTCAGAGATTTATTCGATCTCGTGATTAAGGCTTAATCTACTACAAATCTC 1607
Db 347 LeuAspPheArgAspLeuTyrrAspProArgAspLysAlaLysLeuLeuTyrrAsnAsnLeu 366
Qy 1608 GATGCTTTGGTATTATGAGCTATACCTTAACTGGAAAGAGAGATTAATCAAGATGAC 1667
Db 367 AspAlaPheGlyLysLeuAspTyrrThrLeuThrglyLysValGluAspAsnHisAspAsp 386
Qy 1668 ACCAACCGTATCATTAACGTTTATATGAGAGCAAGCCGAGAGAGAGAAATGCTAGCTAC 1727
Db 387 ThrAsnArgIleIleThrValTyrrMetGlyLysArgProGluGlyGluAsnAlaSerTyrr 406
Qy 1728 CATTTAGCT 1736

Db 407 HistLeuAla 409

RESULT 7

US-07-854-596B-47

Sequence 47, Application US/07854596B

Patent No. 5434073

GENERAL INFORMATION:

APPLICANT: Dawson, Keith M

APPLICANT: Hunter, Michael G

APPLICANT: Czaplowski, Lloyd G

TITLE OF INVENTION: Proteins and nucleic acids

NUMBER OF SEQUENCES: 73

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dr. John J. McDonnell

STREET: Ten South Wacker Drive, Suite 3000

CITY: Chicago

STATE: IL

COUNTRY: USA

ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/854,596B

FILING DATE: 03-JUN-1992

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: McDonnell, John J

REGISTRATION NUMBER: 26,949

REFERENCE/DOCKET NUMBER: 92,337

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-715-1000

TELEFAX: 312-715-1234

TELEX: 910-221-5317

INFORMATION FOR SEQ ID NO: 47:

SEQUENCE CHARACTERISTICS:

LENGTH: 483 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-07-854-596B-47

US-09-940-235-12 (1-2096) x US-07-854-596B-47 (1-483)

QY 588 ATTGTTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACGACCAATGTTGTTT

Db 1 IleaaglyProglutirleuLeuSerProSerValAsnAsnSerGlnLeuValVal

QY 648 AGCGTTGCTGACTGTTGAGGGGAGCAATCAAGACATTAAGTCTTAATTTTGAATC

Db 21 ServaiaaiglyThValaigluThThraaGlnAspIleSerLeuYsrPheGluIle

QY 708 GATTAACATCAAGACCTGCTGATGAGAGAAAGACAGAGAGCTTAAGTCCAAATCA

Db 41 AspleuThrSerArgProIleAsnIleGlyGlyThGlnGlnIleuSerProIlySer

QY 768 AAACATTTGCTACTGATGATGGCCGAGATGTCACATAACTTGAAGAAAGTGACTTA

Db 61 LysPhePheIaThrAspSerGlyAlaMetProIleLeuLeuIleuValAspLeu

QY 828 AAGGTAATTCAGACATTAATGATGCTAAGTCCACAGTAAGACGATTAATTTGAAGTC

Db 887

81 LysAlaIleGlnGlnIleuIleAlaAsnValHisSerAsnAspPheGlyVal

QY 888 ATTGATTTTGACAGGATGCAACCATTAATGATGCAACCGGCAAGCTTACTTGTCTAC

Db 101 IleAspPheIaSerAspAlaThrIleThrAspAsnGlyLysValIlePheIleAsp

QY 948 AAGATGGTTGCTTACCTTCCGACCAACCTGTCCAAAGATTTTCTTAAGCGGACAT

Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnGlnIlePheLeuSerGlyHis

QY 1008 GTGCGCGTTAGACCATTAATAAGAAACCAATACAAACCAAGCAAACTTGTGATGTG

Db 141 ValArgValArgProGlyGlyGlySerProIleGlnAsnGlnAlaLysSerValAspAl

QY 1068 GAATATCTGTACAGTTTACCTTCCCTTAACCTGATGACGATTTTCAGACCGCTTCAA

Db 161 GluThrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys

QY 1128 GATACATAAGCTATTGAAAACATGACTATCGGTGACACATCAATCTCAAGAAATTA

Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnIleuLeu

QY 1188 GCTCAGACCAAGACATTTTAAACAAAACCAAGGCTTATGACATTTTATGACAGTAC

Db 201 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyThrIleGlyGlnArgAsp

QY 1248 TCCCTCATTCGCTCATGACATGACATTTTCCGTACGATTTTACCAATGATGACAG

Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnIle

QY 1308 TTTACTTACCGTGTAAATAATCGGGAACAAGCTTATGATGATCAATTAATAATAATCG

Db 241 PheThrThrValValLysAsnArgGlnGlnAlaThrGlnIleAsnLysSerGlyLeu

QY 1368 AATGAGAAATTAACAACACTGACCTGATCTGTGAGAAATTAATGCTTAAATAAGGG

Db 261 AsnGlnGlnIleAsnAsnThrAspLeuIleSerGlnLysThrValLeuLysGly

QY 1428 GAAAGCCGATGATCCCTTGTGATGCACTGATGCACTTGAACCTGTCACCATTAAT

Db 281 GluLysProLysAspProPheAspArgSerHisLeuLysLeuPheThrIleLysThrVal

QY 1488 GATGTGATACCAAGCAATGTCTTAAAGGAGAGGCTTAAAGCTTAAAGCTTAAAG

Db 301 AspValAsnThrAsnGlnLeuLeuLysSerGlnIleuLeuThrAlaSerGlnArgAsn

QY 1548 TTAGACTTCAGAGATTATACGATCTGTGATGAGCTTAACTTCAACAATCTC

Db 321 LeuAspPheArgAspLeuThrAspProArgAspLysAlaLysLeuLeuThrAsnAsnLeu

QY 1608 GATGCTTTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT

Db 341 AspAlaPheGlyIleLeuSerArgThrLeuThrGlyLysValGlnAspAsnHisAspArg

QY 1668 ACCAAGCGTATCAATACCGTTATATGAGGAGAGAGAGAGAGAGAGAGAGAGAG

Db 361 ThrAsnArgIleIleThrValIleThrValIleThrValIleThrValIleThrValIle

QY 1728 CATTAAGCTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG

Db 381 HisLeuAla-----LysArgLysArgPheThrGlnGlnGlnGlnGlnGlnGlnGln

QY 1788 CATGCTGCTGAGACTTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG

Db 399 TyrLeuArgGlyThrGlnThrProIleProIleProIleProIleProIleProIlePro

QY 1848 ATGATG---GTAAATGCTACT-----TGCCTGGAAGAGGAGGAGGAGGAGGAG

Db 419 ValValValThrAspCysThrGlnSerGlyGlnAsnLeuLysCysGlnGlnGlnGlnGln

QY 1887 CGGATCACTTGCACTTCAAGATTAATGATGATGATGATGATGATGATGATGATGATG

Db 439 -----ValCysGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln

447


```

QY      1947 ATTGAGACACTGAGCAGAAGATATCATGAGAACTCGTCACGTGCATGCGACA   2006
           ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      448 leuglserrapbglyglulubahn-----glnCyval---Thr 459
QY      2007 GCGAACGCCGAGAGAGTGGAAAGTTGTGAGAGCACCTCTGTGCACACATCGAGC   2066
           ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      460 Glylugly-----ThrProlyPrglInserHis 469
QY      2067 GGATCTGGCCCTTCACCGATGTT 2090
Db      470 AenapGlyAepPneGlugluile 477

RESULT 8
US-08-488-940-1
; Sequence 1, Application US/08488940
; Patent No. 5854049
; GENERAL INFORMATION:
; APPLICANT: Reed, Guy L.,
; TITLE OF INVENTION: PLASMIN-RESISTANT STREPTOKINASE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,940
; FILING DATE: 09-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Frazer, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 05433/009001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1194 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-488-940-1

Alignment Scores:
Pred. NO.:          Length:       1194
Score:              1976.50     Matches:        389
Percent Similarity:  97.50%     Conservative:    1
Best local Similarity: 97.25%     Mismatches:      5
Query Match:         53.05%     Indels:          5
DB:                  Gaps:       3

US-09-940-235-12 (1-2096) x US-08-488-940-1 (1-1194)
QY      546 CAGACCACCTGAGAC-----GGATCTGGCCCCCTTCAACCGCATGTTGGATTGCTGA   596
           ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      366 GlnthrnsnbsrSerSeValPrgIlyArgIySerile--GluGlyArgIleAlagly 384
QY      597 CCTGAGTGGCTGCTAGACCCGTCATCTGTCAAACAAGACCAATGATGTTAGCGGTGCT   656
           ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      385 PrglutlrPlenLeuAepArygPröSerValasnhsnrGlnLeuValSerValala 404
QY      657 GGRACGTGTGAGGGAGCAATCAAGACATTAGCTTAATAATTTTTGAANTGATCTACA   716

```

D	405	GlyThrValIGluGlyThrAsnGlnAspIleSerLeuValPhePheGluIleLeuAspLeuThr	424
Q	717	TCAGCACTGCTCATGAGAGAAAGACAGACGAGCTTAAGTCCAAATCAAAACCAATT	776
D	425	SerArgProAlaHISGlyGlySerThrGlnGlnGlyLeuSerProIleSerValPhe	444
Q	777	GCTACTGATNAGGGGGCGAGTGCATATAACTTGGAGAAAGCTGACTTACTTAAGCTATT	836
D	445	AlaThrAspSerGlyAlaMetSerHisIleValLeuGluValAspLeuLeuValIle	464
Q	837	CAAGAACCAATTGATGCTTAACGTCACAGTAAACGACACTCTTGGAGGTCATTGATTT	896
D	465	GlnGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluValIleAspPhe	484
Q	897	GCAACCGATGCACCACTTACTGATGAAACGGCAAGGTCTTACTTGGCTGACAAAGATG	956
D	485	AlaSerAspAlaThrIleThrAspArgAsnGlyValTyrPheAlaAspValAspGly	504
Q	957	TCGGTAACCTTGGCCGACCCACCTGTCCAAGAATTTTGGCTAAGGGGACATGTGGCGTT	101
D	505	SerValThrLeuProThrGlnProValGlnGlnPheLeuLeuSerGlyHisValArgVal	524
Q	1017	AGACCATTAATAAGAAACCAATACAAAACCAAGGAATCTGTTGATGAGTAATTACT	1077
D	525	Arg---TyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspValGluTyrTrp	543
Q	1077	GTACAGTTTACTCCCTTAAACCTCGATGACGATTTCAAGCAACGCTCCAAAGATCTAAG	1137
D	544	ValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLysAspThrLys	563
Q	1137	CTATTGAAAAACATNAGCTATGCGTGACACCATCATCTCAAGAAATTACTAGCTCAGCA	1197
D	564	LeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeuLeuAlaGlnAla	583
Q	1197	CAAGCATTTTAAACAAAACCAACCAAGGCTATACAGTTTATGAACGTCCCTCAATC	1257
D	584	GlnSerIleLeuAsnLysAsnHisProGlyTyrTrpIleTyrGlnArgAspSerIle	603
Q	1257	GTCACTCATGACCAATGACATTTTCCGTCGATTTTACCAATGATCAAGATTTACTTAC	1317
D	604	ValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnIlePheThrTyr	623
Q	1317	CGTGTAAAAATCGGGAACAAGCTTATNAGATCAATAAAAATCTGGTCTGATGAGAA	1377
D	624	ArgValLysAsnArgGlnGlnAlaTyrArgIleAsnLysSerGlyLeuAsnGlnGlu	643
Q	1377	ATTAACCACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGGGAAAAACCG	1437
D	644	IleAsnAsnThrAspLeuIleSerGlyLysTyrTyrValLeuLysIleGlyLysPro	663
Q	1437	TATGATCCCTTGGATCGCAGTCACTTGGAACTGTTCACATCAAAATACCTTATGTCGAT	1497
D	664	TyrAspProPheAspArgSerHisIleuLysLeuPheThrIleLysTyrValAspValAsp	683
Q	1497	ACCAACGAATTGCTAAAAAGTAGCAGGCTTTAACAGCTTACGGAACGTACTTAAAGCTTC	1557
D	684	ThrAsnGlnLeuLeuLysSerGlyGlnLeuLeuThrAlaSerGlyAlaArgAsnLeuAspPhe	703
Q	1557	AGAGATTATACGATCCCTCGATATAGGCTAAACTACTTACACAACTTCGATGCTTTT	1617
D	704	ArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeuAspAlaPhe	723
Q	1617	GGATTATGAGCTAATCCTTAACTGGAAGAGTAGAGATATACAGATGACCAACGCT	1677
D	724	GlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAspThrAsnArg	743
Q	1677	ATCATTAACCGTTTATATGAGGACAGGCAACCGCAAGAGAGAAATGCTAGCTACATTAGCT	1737
D	744	IleIleThrValTyrMetGlyLysArgProGlnGlyGluAsnAlaSerTyrHisLeuAla	763

```

; Sequence 28, Application US/07854596B
; Patent No. 5434073
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith M
; APPLICANT: Hunter, Michael G
; APPLICANT: Czaplowski, Lloyd G
; TITLE OF INVENTION: Proteins and nucleic acids
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. John J. McDonnell
; STREET: Ten South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/854,596B
; FILING DATE: 03-JUN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,337
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 499 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-854-596B-28

```

```

Alignment Scores:
Pred. No.: 4,066-188 Length: 499
Score: 1969.00 Matches: 382
Percent Similarity: 96.00% Conservative: 2
Best Local Similarity: 95.50% Mismatches: 16
Query Match: 52.84% Indels: 0
Gaps: 0
US-09-940-235-12 (1-2096) x US-07-854-596B-28 (1-499)

```

```

QY 537 ACCTTGTGACAGACACATCGAGCGGATCTGCCCTTACCGATGTCGTAATGCTGGA 596
Db 69 Thrilealaserillealalalaysglugluglyvalserleuasparysargillealagly 88
QY 597 CCGTGGGCGCTGCTGACCGGCAATCTGTCACAAAGCAAGCAATGTTGTAACGCTTGT 656
Db 89 ProgluttrpleuasparyproserValaasnsercInleuValalSerValaA 108
QY 657 GGTACTGTTGAGGGAGCAATCAAGACATTAAGTCTTAAATTTTGAATCGATTAACA 716
Db 109 Glythrvalgluglythrshenlnaspilserleuylpshetheglulleaspheuthr 128
QY 717 TCACGACCTGCTCATGAGAGAAAGACAGAGAGGCTTAAGTCAAAATCAAAACCATTT 776
Db 129 SerargProalaHiegllyllystrngluglnglyleuSerrolyserysprophe 148
QY 777 GCTACTGATAGTGGCGCATGTCACATTAACCTTGAAGAGCTGACTTAAAGGCTATT 836
Db 149 AlathraspserglYalamerProhlsylsleuglulysalaspheuleuysalalle 168
QY 837 CAAGAACAATTGATGCTTAAGTCCACAGTAACGAGACACTTGAAGGATTAATTTT 896
Db 169 GlngluglnleuilealasnValHisSerAsnasparyrthegluValilleasphe 188

```

```

QY 897 GCAAGCATGCAACCATTAATGATGAAACGGCAAGGTCTACTTTGTCGACAAAGATGT 956
Db 189 AlaserapalatrillethrAspArgasngllylsvalTyrPhealasparysaglly 208
QY 957 TCGGTAACTTGGCCGACCAACCTGTCCAGAAATTTTGTAAAGGACATGTGCCGTT 1016
Db 209 SerValThrleuProthrInProValglnglnpHeuleuSerGlyHisValArgVal 228
QY 1017 AGACCATTTAAAGAAAAACCAATCAAAACCAAGCAATCTGTTGATGTGGAATTAAT 1076
Db 229 ArgProTyrlysglyysProilleglnasnlnalalysSerValaspValgllyrthr 248
QY 1077 GTACAGTTTACTCCCTTAAACCTGATACGATTCACACAGGCTCAAGATCTAG 1136
Db 249 ValglInpethrProleuasnProaspaspPheargProglYleuylaspThrlys 268
QY 1137 CTATTGAAAACATAGCTATGCGTGACACCATCATCTCAAGATTTACTAGCTCAAGCA 1196
Db 269 leuEulysThrleuAlalleglYaspThrillethrserglnglnleuAlaglnla 288
QY 1197 CAAGCATTTTAAACAAAACACCCAGGCTATACGATTTAAGACGTGACTCTCAATC 1256
Db 289 GlnserilleuEuanlyThrHisProglYrThrilleTyrGluArgAspserille 308
QY 1257 GTCACTGATGACATGACATTTTCCGTACGATTTTCAACATGATCAAGATTTACTAC 1316
Db 309 ValthrHisapnaaplelPheargThrilleuPrometaspGlnpInpethrlyr 328
QY 1317 CGTGTAAAAATCGGGAACAAGCTTATAGATCAATAAATAATCGTCTGAATGAAGA 1376
Db 329 HisValysasnaggluglnalalyrGluileasnlylsysseglYleuasnlglnu 348
QY 1377 ATAAACAACCTGACTGATCTTGAGAAATATTAACGCTTAAAAAAAGGAAAAAGCG 1436
Db 349 IleasnenthraPleulleleserGluysrlyrYrValleuylsYglYllysp 368
QY 1437 TATGATCCCTTTGATCCGACGTCATTTGAAATCTGTTCACCAATATACGTTGATGCAT 1496
Db 369 TyraspProPheaspArgserHisleuylsleuPethrilleTyrValaspValasn 388
QY 1497 ACCAACAATTTGCTAAAGAGAGAGAGCTCTTAACAGCTAGCAAGCACTTGAAGCTTC 1556
Db 389 ThrasnlnleuEulysserglnglnleuEuthrIaserglulysasnleuasphe 408
QY 1557 AGACATTTATACGATCTCGTATAGGCTTAACACTCTCAACAATCTGATGCTTTT 1616
Db 409 ArgaspLeuTyraspProalaspLyalalyseuEuthrIasnleuasphe 428
QY 1617 GGTATTATGACTATACCTTAAGTGAAGAAAGTAAAGATTAATCAAGATGACCAACCGT 1676
Db 429 GlyllewetsapryrThrleuthrGlylyseValgllyasnHiaspAspThrAsnArg 448
QY 1677 ATCATTAACGTTTATATAGGCAAGCGAACCAAGAGAAATGCTAGTACATTTAGCT 1736
Db 449 llellethrVallyrmetGlylsArgProglnglgluasnAlaserTyrHisleuAla 468

```

RESULT 10

```

US-08-488-940-3
; Sequence 3, Application US/08488940
; Patent No. 5834049
; GENERAL INFORMATION:
; APPLICANT: Reed, Guy L.
; TITLE OF INVENTION: PLASMIN-RESISTANT STREPTOKINASE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:

```

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,940
FILING DATE: 09-JUN-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Frazer, Janis K.
REGISTRATION NUMBER: 34,819
REFERENCE/DOCKET NUMBER: 05433/009001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 813 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-488-940-3

Alignment Scores:
Pred. No.: 9,456-188 length: 813
Score: 1966.50 Matches: 382
Percent Similarity: 99.74% Conservative: 0
Best Local Similarity: 99.74% Mismatches: 0
Query Match: 52.78% Indels: 1
DB: 1 Gaps: 1

US-09-940-235-12 (1-2096) x US-08-488-940-3 (1-813)

QY 588 ATTGCTGACCTGAGTGGCTGCTAGACCGTCATCTGTCAACACAGCAATGTTGTT 647
Db 1 IleaaglyproglutripheuleuaspaproserValashnserserleuVal 20
QY 648 AGCGTGTGCTGCTGCTGAGGGAAGAAATCAAGACATTAGTCTTAATTTTGAATC 707
Db 21 ServaIaaglythrValgluIglYthraangIaapIleSerleuysphegnIuIle 40
QY 708 GATCTAACATCACGACCTGCTCATGTGAGAGAAAGACGACCAAGCTTAAGTCCAAATCA 767
Db 41 AspleuttrSerArgproIahIsglyIystrhgluIglYleuSerProIysSer 60
QY 768 AAACCATTTGCTAGTATGAGCGCGAGTGTCAACATTAAGTGAAGAGCTGACTACTA 827
Db 61 LysproPhealathAspSerIyalaMetSerHlsleuIglulysalaaspleu 80
QY 828 AAGCTATTCAAGAACCAATTGATCGCTAACGTCACAGTACGACACTACTTGAAGTC 887
Db 81 LysalaIleIgluIgluIleuIlealaenValHisSerAsnaspIyrrphegluVal 100
QY 888 ATTGATTTTGCAGCGAGTACCAACATTAAGTGAAGAGCGAGGCTTACTTGTCTGAC 947
Db 101 IleaspPhealaseAspalaIthrIlethraAspArgsnIyIysValTyrrPhealaasp 120
QY 948 AAAGATGCTGCGTAACTTGGCCGACCCAACTGTCACAGAAATTTTGTGAAGCGGACAT 1007
Db 121 LysaspGlySerValThrleuprothrIgmprovalIgluIgluIleuSerGlyHis 140
QY 1008 GTGCGCGTTAGACCATATTAAGAAAAACCAATACAAACCAAGCAAACTGTTGATGTG 1067
Db 141 ValargValarg---TyrIleGlyIystrIleIgluIgluIlealysSerValaspVal 159
QY 1068 GAATTAATCTGACATGTTACTCCCTTAAACCTGATGACGATTTTGACCAAGGCTCGAA 1127
Db 160 GluTyrrThrValIgluIgluIleuIleuIleuIleuIleuIleuIleuIleuIleu 179
QY 1128 GATATAAGCATTTGAAACATAGCTATCGGTGACACCATCAACATCTCAAGAAATCACTA 1187

Db 180 AspThrIysleuIleuIystrIleuAlaIleGlyAspThrIleThrSerGlnIleuIleu 199
QY 1188 GCTCAAGCAACAAGCAATTTTAAACAACACCCAGGCTATACGATTTATGACAGTAC 1247
Db 200 AlaGlnIaIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 219
QY 1248 TCCTCAATCGTCACTGACATGACATGACATTTTCGTAAGATTTTACCAAGATCAAGAG 1307
Db 220 SerSerIleValIthrIshAspAsnaspIlePheargThrIleleuprometAspGlnIu 239
QY 1308 TTTACTTACCGTGTAAATAATCGGAAACAAGCTTATGATCAATAAAAAATCTGCTCTG 1367
Db 240 PheThrTyrrArgValIyasaenArgIgluIgluIleuIleuIleuIleuIleuIleuIleu 259
QY 1368 AATGAAGAAATAACAACCTGACCTGATCTGAGAAATATACGCTTAAATAAAAGG 1427
Db 260 AsnIgluIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 279
QY 1428 GAAAGCCGATATGATCCCTTGTATGCGAGTCACTGAACTGTTCACCATCAATACGTT 1487
Db 280 GluIystrProIyrrAspProPheaspArgSerHlsleuIystrleuThrIleIystrVal 299
QY 1488 GATGTCGATACCAAGCAATGCTAAAGTACAGCTCTTAAACGCTTACCGAAAGTAC 1547
Db 300 AspValAspThrangIleuIleuIystrGluIgluIleuIleuIleuIleuIleuIleuIleu 319
QY 1548 TTAGACTTCGAGATTTATACGATCTCTCGATTAAGGCTTAACCTCTACCAATCTC 1607
Db 320 LeuaspPheargAspleuTyrrAspProIyrrAspIyalaIystrleuIyrrAsnAsnIleu 339
QY 1608 GATGCTTTGATTTATGAGCTATACCTTACCTTACGAAAGATGAGATTAATCAGATGAC 1667
Db 340 AspAlaPheGlyIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 359
QY 1668 ACCAACCTATCATTAACCTTTTATATGCGCAACGACCCGACAGAGAGAAATGCTAGTAC 1727
Db 360 ThrAsnArgIleIleThrValTyrrMetGlyIyArgProIgluIgluIleuAlaSerTyrr 379
QY 1728 CATTAGCT 1736
Db 380 HisleuAla 382

RESULT 11
US-07-854-596B-26
Sequence 26, Application US/07854596B
Patent No. 5434073
GENERAL INFORMATION:
APPLICANT: Dawson, Keith M
APPLICANT: Hunter, Michael G
APPLICANT: Czaplowski, Lloyd G
TITLE OF INVENTION: Proteins and nucleic acids
NUMBER OF SEQUENCES: 73
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dr. John J. McDonnell
STREET: Ten South Wacker Drive, Suite 3000
CITY: Chicago
STATE: IL
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/854,596B
FILING DATE: 03-JUN-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: McDonnell, John J
REGISTRATION NUMBER: 26,949
REFERENCE/DOCKET NUMBER: 92,337
TELECOMMUNICATION INFORMATION:

: TELEPHONE: 312-715-1000
 : TELEFAX: 312-715-1234
 : TELERX: 910-221-5317
 : INFORMATION FOR SEQ ID NO: 26:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 415 amino acids
 : TYPE: amino acid
 : TOPOLOGY: linear
 : MOLECULE TYPE: protein
 : US-07-854-596B-26

 Alignment Scores:
 Pred. No.: 4,64e-187 Length: 415
 Score: 1958.00 Matches: 378
 Percent Similarity: 98.96% Conservative: 1
 Best Local Similarity: 98.69% Mismatches: 4
 Query Match: 52.55% Indels: 0
 DB: 1 Gaps: 0

 US-09-940-235-12 (1-2096) x US-07-854-596B-26 (1-415)

QY 588 ATTGCTGACCTGAGTGGCTGTAGACGCTCATCTGTCAACAGCAATTTGCTTTT 647
 Db 2 TTTAAGTGTGCTGCTGTGAGGAGGAGCAATCAAGCTTAAATTTTGGAAATC 21
 QY 648 AGCGTGTGCTGCTGTGAGGAGGAGCAATCAAGCTTAAATTTTGGAAATC 707
 Db 22 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuYsheheGlu 41
 QY 708 GATCTTAACATCAAGCCTGCTCATGAGAGAAAGACAGCAAGCTTAAAGTCCAAATCA 767
 Db 42 AspLeuThrSerArgProAlaHisGlyGlyThrGlnGlnGlyLeuSerProlySer 61
 QY 768 AAATCATTGCTGCTGATGAGTGGCGGCTGTCACATTAATTTGAGAAAGCTGATCTTA 827
 Db 62 LysProPheAlaThrAspSerGlyAlaMetProHISLysLeuGlnLysAlaAspLeu 81
 QY 828 AAGCTATTCAAGAAACAATTGATGCTGCTAAGCTCAAGCAAGTCAAGTCTTGAAGTC 887
 Db 82 LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspPyrRheGluVal 101
 QY 888 ATTGATTTTGCAGGAGTGAACAACATTAATGTCGAAACGCAAGTCTTACTTGTGAC 947
 Db 102 IleAspPheAlaSerAspAlaThrIleThrAspArgGlnGlyLysValTyrRheAlaAsp 121
 QY 948 AAAAGATGTTGGTGAACCTTGGCCGACCACTGTGTCAGAAATTTTGTAAAGCGACAT 1007
 Db 122 LysAspGlySerValThrLeuProThrGlnProValGlnGlnPheLeuSerGlyHis 141
 QY 1008 GTGGCGGCTTAAACCATTAAGAAACCAATTAACAAACCAAGCAAGTCTTGAATGTC 1067
 Db 142 ValArgValAlaGlyProTyrLysGlnLysProIleGlnAsnGlnAlaLysSerValAspVal 161
 QY 1068 GAATATCTGTAAGTTTACTCCCTTAAACCTGATGACATTTGAGACAGAGTCTCAAA 1127
 Db 162 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProIleLys 181
 QY 1128 GATATCAAGTATTGAAACCACTAGCTTACGCTGACACCATCAATCTTCAAGAAATTA 1187
 Db 182 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 201
 QY 1188 GCTCAAGCAAGCAATTTTAAACAAACCAAGGCTTATGCAATTTTGAAGTGCAC 1247
 Db 202 AlaGlnAlaGlnIleLeuAsnLysThrHisProGlyTyrThrIleTyrGlnArgAsp 221
 QY 1248 TCCCAATGCTGATGATGACATGATGATTTTCCGTAGATTTTCAATGATGATCAAGAG 1307
 Db 222 SerSerIleValIleHisAspAsnAspIlePheArgThrIleLeuProMetCaspGlnGlu 241
 QY 1308 TTATCTTAACCGTGTAAATAATCGGGAACAAGCTTATAGACATTAATAATAATCTGATCTG 1367
 Db 242 PheThrTyrHisValLysAsnArgGlnGlnAlaTyrGlnIleAsnLysLysSerGlyLeu 261

QY 1368 AATGAAGAAATAAACAACACTGACCTGATCTTGAGAAATATTACGCTTAAAAAGG 1427
 Db 262 AsnGlnGluIleLeuAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysGly 281
 QY 1428 GAAAAAGCCGTATGATCCCTTTGATGACATCTTGAACACTGTCACCATCAATACGTT 1487
 Db 282 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleTyrVal 301
 QY 1488 GATGCTGATCAACGAATTTGCTAAAGAGAGAGCTTAAAGCTTAAAGCTTAAAGCTTAA 1547
 Db 302 AspValAsnThrAsnGlnLeuLysSerGlnGlnLeuLysTrpAlaSerGlnArgAsn 321
 QY 1548 TTAGACTTCAAGATTTTATAGATCTGATGATGAGCAAGCAAGCAAGCAATCTTAACTTC 1607
 Db 322 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLysTyrAsnAsnLeu 341
 QY 1608 GATGCTTTGATTTATGACATATACCTTAACTGAAAAAGTAAAGATTAATCAGATGAC 1667
 Db 342 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 361
 QY 1668 ACCAAGCTATCATTAACCGTTTATATGAGGAGCAAGCAAGCAAGCAAGCAATCTTAACTTC 1727
 Db 362 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGlnGlnGluAsnAlaSerTyr 381
 QY 1728 CATTTAGCT 1736
 Db 382 HisLeuAla 384

 RESULT 12
 US-07-854-596B-19
 : Sequence 19, Application US/07854596B
 : Patent No. 5434073
 : GENERAL INFORMATION:
 : APPLICANT: Dawson, Keith M
 : APPLICANT: Hunter, Michael G
 : APPLICANT: Czaplowski, Lloyd G
 : TITLE OF INVENTION: Proteins and nucleic acids
 : NUMBER OF SEQUENCES: 73
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Dr. John J. McDonnell
 : STREET: Ten South Wacker Drive, Suite 3000
 : CITY: Chicago
 : STATE: IL
 : COUNTRY: USA
 : ZIP: 60606
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: PatentIn Releasee #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/07/854,596B
 : FILING DATE: 03-JUN-1992
 : CLASSIFICATION: 435
 : ATTORNEY/AGENT INFORMATION:
 : NAME: McDonnell, John J
 : REGISTRATION NUMBER: 26,949
 : REFERENCE/DOCKET NUMBER: 92,337
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 312-715-1000
 : TELEFAX: 312-715-1234
 : TELERX: 910-221-5317
 : INFORMATION FOR SEQ ID NO: 19:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 435 amino acids
 : TYPE: amino acid
 : TOPOLOGY: linear
 : MOLECULE TYPE: protein
 : US-07-854-596B-19

 Alignment Scores:
 Pred. No.: 4,76e-187 Length: 435
 Score: 1958.00 Matches: 378

Percent Similarity: 98.96%
 Best Local Similarity: 98.69%
 Query Match: 52.55%
 DB: 1
 Conservative: 1
 Mismatches: 4
 Indels: 0
 Gaps: 0

US-09-940-235-12 (1-2096) x US-07-854-596B-19 (1-435)

```

OY 588 ATTCGTCGACCTGAGGCGCTGACGCGCTGATCTGTCAACAAGCAAGCAATGTTGTT 647
DB 22 IleaAGlyProGluThrLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 41
OY 648 AGCGTGTGCTACTGTTGAGGGGAGCAATCAAGCATTAGTCTTAATTTTGAATC 707
DB 42 SerValaIaGlyThrValGluGlyThrAsnGlnAspIleSerLeuysPheNegIuIle 61
OY 708 GATCTAACATCCGACCTGCTCTGATGAGGAAAGACAGACAGCGCTTAAGTCCAAATCA 767
DB 62 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 81
OY 768 AAACCATTTGCTACTGATGTCGCGGATGTCACATAAATCTTGAGAACTGATCTACTA 827
DB 82 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGlnLysAlaAspLeuLeu 101
OY 828 AAGGCTATTCAAGAACAAATTTGATCGCTAACGTCACAGTAACGACACTTGTGAGTC 887
DB 102 LysAlaIleGlnGlnGlnLeuLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 121
OY 888 ATGATTTTGCAGCGAGTCAACCATTTACTGATCGAACGCGAGGTCATCTTGTGTAC 947
DB 122 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyValValTyrPheAlaAsp 141
OY 948 AAAAGATGTTGGTGAACCTTGGCGAACCCGACCTGTCCAAAGATTTTGTAGCGGACAT 1007
DB 142 LysAspGlySerValThrLeuProThrGlnProValGlnGlnPheLeuLeuSerGlyHis 161
OY 1008 GTGCGCGTTAGACCATATTAAGAAAAACCAATACAAACCAAGCGCAATCTGTGATGTC 1067
DB 162 ValaGValaIaArgProTyrIleGlnGlyLysProIleGlnAsnGlnAlaLysSerValaAspVal 181
OY 1068 GAATTAAGTGTACAGTTTATCTCCCTTAAACCTTGATGACGATTTTGACGAGCTCTCAA 1127
DB 182 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 201
OY 1128 GATCTAAGCTATTGAAAAACACTAGGATTCGGTGACACCATCACTCAAGAAATTACTA 1187
DB 202 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeuLeu 221
OY 1188 GCTCAAGACAAAGCATTTTAAACAAAAACACCCGAGGCTATACGATTTATGAAGCTGAC 1247
DB 222 AlaGlnAlaGlnSerIleLeuAsnLysThrHisAspGlyTyrThrIleTyrGlnAlaArgAsp 241
OY 1248 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTCAATGATCAAGAG 1307
DB 242 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGln 261
OY 1308 TTTACTTACCGGTTTAAAAATGGGAAACAAGCTTATAGATCAATAAAAATTTGGTCTG 1367
DB 262 PheThrTyrHisValaLysAsnAspArgGlnGlnAlaTyrGlnIleAsnLysLysSerGlyLeu 281
OY 1368 AATGAAGAAATTAACAACAACATGACCTGATCTCGAGAAATATTAAGCTTAAAAAAGG 1427
DB 282 AsnGlnGlnIleAsnAsnThrAspLeuIleSerGlnLysTyrTyrValLeuLysLysGly 301
OY 1428 GAAAAAGCGTATGATCCCTTTGATCGACATCTTGAACCTGTTCACCAATACGTT 1487
DB 302 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 321
OY 1488 GATGTCGATPACCAAGAAATGCTTAATAAAGTGAAGACGCTTTAAACAGCTACGGAAGTAA 1547
DB 322 AspValaAsnThrAsnGlnLeuLeuLysSerGlnGlnLeuLeuThrAlaSerGlnAlaGln 341
OY 1548 TTAGCTTCAGAGATTTTATACGATCTCGTGAATAGGCTAACTACTACAAACATCTTC 1607

```

```

DB 342 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 361
OY 1608 GATGCTTTGGTATTTATGACCTATACCTTAACTGGAACGTAATCAAGCAAGTAC 1667
DB 362 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGlnAspAsnHisAspAsp 381
OY 1668 ACCAAGCTATCAATACCGTTTATATGGGCAAGCAAGCAAGGAGAGAAATGCTAGCTAC 1727
DB 382 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGlnGlnAlaSerTyr 401
OY 1728 CATTAGCT 1736
DB 402 HisLeuAla 404

RESULT 13
US-07-854-596B-15
; Sequence 15, Application US/07854596B
; Patent No. 5434073
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith M
; APPLICANT: Hunter, Michael G
; APPLICANT: Czaplowski, Lloyd G
; TITLE OF INVENTION: Proteins and nucleic acids
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Dr. John J. McDonnell
; STREET: Ten South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/854,596B
; FILING DATE: 03-JUN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,337
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ. ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 440 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-854-596B-15

Alignment Scores:
Pred. No.: 4,79e-187 Length: 440
Score: 1958.00 Matches: 378
Percent Similarity: 98.96% Conservative: 1
Best Local Similarity: 98.69% Mismatches: 4
Query Match: 52.55% Indels: 0
DB: 1 Gaps: 0

US-09-940-235-12 (1-2096) x US-07-854-596B-15 (1-440)
OY 588 ATTCGTCGACCTGAGGCGCTGACGCGCTGATCTGTCAACAAGCAAGCAATGTTGTT 647
DB 22 IleaAGlyProGluThrLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 46
OY 648 AGCGTGTGCTACTGTTGAGGGGAGCAATCAAGCATTAGTCTTAATTTTGAATC 707
DB 42 SerValaIaGlyThrValGluGlyThrAsnGlnAspIleSerLeuysPheNegIuIle 66

```

QY 708 GATCTAATCAGCAGCTGCTGATGAGAAAGACAGAGCGCTTAAGTCCAAATCA 767
 DB 67 AspLeuThrSerArgProAlaHisGlyGlySerThrGluGlnGlyLeuSerProLysSer 86
 QY 768 AAACCTTTGCTACTGATAGTGGCGGCGCATGTCATATACTTGAGAAAGCTGACTTACTA 827
 DB 87 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGlyValAlaAspLeuLeu 106
 QY 828 AAGGCTATTGCAAGAACATGATGCTAGCTGACAGTAAAGAGTACTTGGAGGCTC 887
 DB 107 LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspArgTyrPheGluVal 126
 QY 888 ATTGATTTTGCAGACGATGCAACCATTTAGTCAAGAACGGCAAGTCTTACTTGTGAC 947
 DB 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146
 QY 948 AAAGATGGTTCGGTAACTTGGCCGACCCAACTGTGCCAAGAAATTTTGTAAAGCGACAT 1007
 DB 147 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 166
 QY 1008 GGTGGGTTAGACCATTAAGAAACCAATACCAAAACCAAGGGAATCTGTGATGTG 1067
 DB 167 ValArgValArgProTyrLysGlyLysProIleGlnAsnGlnAlaLysSerValAspVal 186
 QY 1068 GAATATACGTGTACGTTTACCTCCCTTAAACCTGATGACGATTCAGACCGAGTCTCAA 1127
 DB 187 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheAlaArgProGlyLysLys 206
 QY 1128 GATCTAAGCTATTGAAACACTAGCTAGTGGTGACACATCATCATCTCAAGATTACTA 1187
 DB 207 AspThrLysLeuLeuLysThrIleAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 226
 QY 1188 GCTCAGACCAAGCATTTTAAACAAACACACCCAGGCTTATGAGATTATGAACGTGAC 1247
 DB 227 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProLysTyrThrIleTyrGluArgAsp 246
 QY 1248 TCCCTAATGCTCATGATGACATGACATTTTCCGTACGATTTTAAAGGATGACAG 1307
 DB 247 SerSerIleValAlaThrHisAspAsnAspIlePheThrGlnThrLeuProMetAspGlnGlu 266
 QY 1308 TTCTACTACCGTGTAAATTCGGGAAACAGCTTATGATCAATATAAAATCGTGTG 1367
 DB 267 PheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsnLysLysSerGlyLeu 286
 QY 1368 AATGAAGAATAAACAACACTGACCTGATGCTGAGAAATATTACGCTCTTAAAAAGG 1427
 DB 287 AsnGluGluIleAsnAsnThrAspLeuIleSerGlyLysTyrTyrValLeuLysLysGly 306
 QY 1428 GAAAGCCGATAGATCCCTTGTATCCAGTCACTTGAAACTGTTCACCAATAGT 1487
 DB 307 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326
 QY 1488 GATGTCGATACCAACGAATTGTCTAAAGAGTACAGCTCTTAACAGCTAGCAACGTAAC 1547
 DB 327 AspValAsnThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGlyLysAsn 346
 QY 1548 TTAGACTTCAGAGATTATACGATCTCGTATAGGCTAAAGCTACTTCAACAATATTC 1607
 DB 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
 QY 1608 GATGCTTTTGGTATTATGACATATACCTTAACTGGAAGAGTAATACAGATGAC 1667
 DB 367 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 386
 QY 1668 ACCAACCGTATGATACCGTTTATATAGGGCAAGCAACCGCAGAGAGAAATGCTAGTAC 1727
 DB 387 ThrAsnAspGlyIleThrValLysMetGlyLysArgProGluGluAsnAlaSerTyr 406
 QY 1728 CATTTAGCT 1736
 DB 407 HisLeuAla 409

RESULT 14
 US-07-854-596B-35
 : Sequence 35, Application US/07854596B
 : Patent No.5434073
 : GENERAL INFORMATION:
 : APPLICANT: Dawson, Keith M
 : APPLICANT: Hunter, Michael G
 : APPLICANT: Czaplowski, Lloyd G
 : TITLE OF INVENTION: Proteins and nucleic acids
 : NUMBER OF SEQUENCES: 73
 : CORRESPONDENCE ADDRESS:
 : ADDRESS: Dr. John J. McDonnell
 : STREET: Ten South Wacker Drive, Suite 3000
 : CITY: Chicago
 : STATE: IL
 : COUNTRY: USA
 : ZIP: 60606
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: PatentIn Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/07/854,596B
 : FILING DATE: 03-JUN-1992
 : CLASSIFICATION: 435
 : ATTORNEY/AGENT INFORMATION:
 : NAME: McDonnell, John J
 : REGISTRATION NUMBER: 26,949
 : REFERENCE/DOCKET NUMBER: 92,337
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 312-715-1000
 : TELEFAX: 312-715-1234
 : TELEX: 910-221-5317
 : INFORMATION FOR SEQ ID NO: 35:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 859 amino acids
 : TYPE: amino acid
 : TOPOLOGY: linear
 : MOLECULE TYPE: protein
 : US-07-854-596B-35
 : Alignment Scores:
 : Pred. No.: 6.92e-187 Length: 859
 : Score: 1958.00 Matches: 378
 : Percent Similarity: 98.96% Conservative: 1
 : Best Local Similarity: 98.69% Mismatches: 4
 : Query Match: 52.55% Indels: 0
 : DB: 1 Gaps: 0
 : US-09-940-235-12 (1-2096) x US-07-854-596B-35 (1-859)
 : QY 588 ATTGCTGACCTGAGTGGCTGCTGATGACCGTGCATCTGTCAACACAGCAATGGTGT 647
 : DB 22 LLeAlaGlyProGluThrPheLeuAspArgProSerValAsnAsnSerGlnLeuValAl 41
 : QY 648 AGCGTTGCTGTACTGTGAGGGGACGATCAACAGATTAGCTTAAATTTTGAATC 707
 : DB 42 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 61
 : QY 708 GATTTAATCATCAGACCTGCTCATGAGGAAAGACAGACGAAAGCTTAAGTCCAAATCA 767
 : DB 62 AspLeuThrSerArgProAlaHisGlyGlySerThrGluGlnGlyLeuSerProLysSer 81
 : QY 768 AAACCTTTGCTACTGATAGTGGCGGCGCATGTCATATACTTGAGAAAGCTGACTTACTA 827
 : DB 82 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGlyValAlaAspLeuLeu 101
 : QY 828 AAGGCTATTGCAAGAACATGATGCTAGCTGACAGTAAAGAGTACTTGGAGGCTC 887
 : DB 102 LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspArgTyrPheGluVal 121
 : QY 888 ATTGATTTTGCAGACGATGCAACCATTTAGTCAAGAACGGCAAGTCTTACTTGTGAC 947

```

Db      122  ILeaSPheAlaSeSerAspAlaThrIleThrAspArgAsnGlyLeuValTyrPheAlaAsp 141
Qy      948  AAAGATGGTTCGGTAACTTGGCCGACCCAACTGTCGAAGATTTTGTACGGCAAT 1007
Db      142  LysaSPgIySeValThrIeuProThrgInProValGInGluPheIeuSeSerGlyHis 161
Qy      1008  GTGGCGGTAGACATATTAAGAAAAACAAATACAAACCAAGCGCAATCTGTGATGG 1067
Db      162  ValArgValArgProTyrIlySeGluYsProIleGlnAsnGlnAlaYsSerValAspVal 181
Qy      1068  GAATATCTGTACAGTTTACTCCCTTAAACCTGATGACATTTTCAGACAGGTCCTCAA 1127
Db      182  GluTyrThrValGlnPheThrProIeuAsnProAspAspPheArgProGlyLeuYs 201
Qy      1128  GATATCAAGCTATTGAAAACACTAGCTATGGTGACACATCAATCTCAAGATTAATA 1187
Db      202  AspThrIlySeIeuIlyeThrIeuAlaIleGlyAspThrIleThSerGInGluLeu 221
Qy      1188  GCTCAAGACAAAGATTTTAAACAAACCAACCGGTATACGATTTATGACGCTGAC 1247
Db      222  AlaGlnAlaGInSerIleIeuAsnIySthrHisProGlyTyrThrIleTyrGluArgAsp 241
Qy      1248  TCCCTCAATCGTCACTCATGACACATGACATTTTCCGTAAGATTTTACCAATGATCAAG 1307
Db      242  SerSerIleValThrHisAspAsnAspIlePheArgThrIleIeuProIeuAspGInGlu 261
Qy      1308  TTTACTTACCGTGTAAATCGGGAACAAGCTTATAGATCAATTAATAAATCTGCTCG 1367
Db      262  PheThrTyrHisValIyAsnArgGInGlnAlaTyrGluIleAsnIlySeSerGlyLeu 281
Qy      1368  AATGAAGAAATTAACAACACTGACCTGATCTGTGAGAAATTTAGCTCTTAAAAAGG 1427
Db      282  AsnGluGluIleAsnAsnThrAspIeuIleSerGluYsTyrTyrValIeuYsIyS 301
Qy      1428  GAAAAGCCGTATGATCCCTTATGACGATGCACTTGAACCTGTGACATCAATACGTT 1487
Db      302  GluIlySProTyrAspProPheAspArgSerHisIleuYsIeuPheThrIleIyS 321
Qy      1488  GATGTCGATACCAAGAAATGCTTAAAGAGAGCTCTTAAACAGCTACCGAACGTAAC 1547
Db      322  AspValAsnThrAsnGluIleuIeuIySeSerGInGluIleuThrIleAspGluArgAsn 341
Qy      1548  TTAAGCTTCAGAGATTTTATGATCTCTGATTAAGCTTAACTACTTCAACAATCTTC 1607
Db      342  LeuAspPheArgAspIeuYsAspProArgAspIySAlaIySeIeuIySAsnAsnIeu 361
Qy      1608  GATGCTTTGGATATGAGCATATACCTTAACCTGAGAAAGTAAAGGATATACAGATGAC 1667
Db      362  AspAlaPheGlyIleMetAspTyrThrIeuIySValGluAspAsnHisAspAsp 381
Qy      1668  ACCAACCGTATCATTAACCGTTTATATGGCAAGCAACCCGAAGAGAGATGCTACTAC 1727
Db      382  ThrAsnArgIleIleThrValTyrMetGlyYsArgProGluGlyIyGluAsnHisAsp 401
Qy      1728  CATTTAGCT 1736
Db      402  HisIeuAla 404

```

RESULT 15

```

US-08-488-940-18
; Sequence 18, Application US/08488940
; Patent No. 5854049
; GENERAL INFORMATION:
; APPLICANT: Reed, Guy L.
; TITLE OF INVENTION: PLASMIN-RESISTANT STREPTOKINASE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESS: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA

```

```

; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,940
; FILING DATE: 09-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 05433/009001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1194 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-488-940-18

```

Alignment Scores:

```

Pred. No.: 9,34e-186 Length: 1194
Score: 1947.50 Matches: 397
Percent Similarity: 87.12% Conservative: 9
Best Local Similarity: 85.19% Mismatches: 50
Query Match: 52.27% Indels: 10
DB: 1 Gaps: 5

```

US-09-940-235-12 (1-2096) x US-08-488-940-18 (1-1194)

```

Qy      546  CAGACCAATTCAGAC-----GGATCTGCGCCCTTACCCGATGTTGATTTGCTGGA 536
Db      366  GlnThrAsnSerSerSerValProGlyArgGlySerIle---GluGlyArgIleAlaGly 384
Qy      597  CCGATGGCGTGTACACGCTCATCTGTCGAACAGCAAGCAATGGTTAGCGGTGCT 656
Db      385  ProGluTrpIeuIeuAspAlaProSerValAsnAsnSerIleuValIseValAla 404
Qy      657  GGTACTGTTGAGGGGAGCAATCAAGCATTAAGTCTTAATTTTGAATCGATCTAACA 716
Db      405  GlyThrValGluGluIyThrAsnGlnAspIleSerIeuAlaPheheGluIleAspLeuThr 424
Qy      717  TCAAGACCTGCTCATGAGGAAAGACAGACAGAGCTTAAGTCCAAAATCAAACTTT 776
Db      425  SerAlaProAlaHisGlyGlyAlaThrGluGlnGlyLeuSerProAlaSerIySProPhe 444
Qy      777  GCTACGATAGTGGCGGAGTGCACATTAACCTTGAAGACGTGACTTAAGGCTATT 836
Db      445  AlaThrAspSerIyAlaMetSerHisIySLeuGluYsAlaAspIeuIySAlaIle 464
Qy      837  CAAGAACAAATTGATCCGTAACGTCAACAGTAAAGACGACTTGAAGTCAATGATTTT 896
Db      465  GInGluGlnIleuIleAlaAsnValHisSerAsnAspAspTyrThrIeuValIleAspPhe 484
Qy      485  AlaSerAspAlaThrIleThrAspArgAsnGlyLeuValTyrPheAlaAspIySAspGly 504
Db      957  TCGGTACCTTGGCGGACCACTGTCGAAGATTTTGTAGCGGACATGGCGGCTT 1016
Qy      505  SerValThrIeuProThrgInProValGInGluPheIeuSeSerGlyHisValArgVal 524
Db      1017  AGACCATATTAAGAAAAACCAATACAAACCAAGCGCAATCTGTGATGGAATATACT 1076
Qy      525  Arg---TyrIySeGluYsProIleGlnAsnGlnAlaIySeSerValaAspValGluTyrThr 543

```

```
OY 1077 GTACGTTTACTCCCTTAAACCTGATGACGATTTCAGACCAAGGCTCAAGAATACTAG 1136
|||
Db 544 ValGlnPheThrProLeuSnpProAspAspPheArgProGlyLeuLysAspThrLys 563
|||
OY 1137 CTATTGAAAAACCTAGCTACCTCGGTGACACCATGACATCTCAAGAAATTACTAGCTCAAGCA 1196
|||
Db 564 LeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnLeuLeuAlaGlnAla 583
|||
OY 1197 CAAGCATTTTAAACAACCAACCAAGGCTATACGATTATGAAACGTGACCTCGCAATC 1256
|||
Db 584 GlnSerIleLeuSnpLysSnpHisProGlyThrIleLeuProMetAspGlnIlePheThrTyr 603
|||
OY 1257 GTCACTCATGACATGACATTTTCGTACGATTTTACCAATGGAATCAAGATTACTTAC 1316
|||
Db 604 ValThrHisAspSnpAspIlePheArgThrIleLeuProMetAspGlnIlePheThrTyr 623
|||
OY 1317 CGTGTAAATAATCGGGAACAAGCTTATAGATCAATTAATAATCTGGTCTGAATGAAGAA 1376
|||
Db 624 ArgValLysAsnArgLysGlnAlaTyrArgIleAsnLysSerGlyLeuAsnGlnGlu 643
|||
OY 1377 ATAAACAACACTGACCTGATCTCTGAGAAATATTACGCTTCTTAAAAAGGGGAAAAGCCG 1436
|||
Db 644 IleAsnAspThrAspLeuLysSerGlnLysTyrTyrValLeuLysLysGlyLysPro 663
|||
OY 1437 TATGATCCCTTGATGACAGTCACTTGAACCTGTTCAACATCAATACGTTGATGTGAT 1496
|||
Db 664 TyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrValAspValAsp 683
|||
OY 1497 ACCAACGAATTGCTAAAAAGTGAAGCAAGCTTTAAACAGCTAGCGAACGTAACTTAGACTTC 1556
|||
Db 684 ThrAsnGlnLeuLeuLysSerGlnLeuLeuThrAlaSerGlnArgAsnLeuAspPhe 703
|||
OY 1557 AGAGATTTTATAGATCTCTGCTGTAAGGCTAACTACTACCAACAATCTCGATGCTTTT 1616
|||
Db 704 ArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeuAspAlaPhe 723
|||
OY 1617 GGTATTATGACTATACCTTAATCTGAAAAGTAGAGATATATCAGATGACCAACCGT 1676
|||
Db 724 GlyIleMetAspTyrThrLeuThrGlyLysValGlnAspAsnHisAspAspThrAsnArg 743
|||
OY 1677 ATCATTAACCGTTTATATGGGCAAGCGAACCGGAAGAGAGATGCTAGCTTACATTTAGCT 1736
|||
Db 744 IleIleThrValTyrMetGlyLysArgProGlnGlyLysAsnAlaSerTyrHisLeuAla 763
|||
OY 1737 GGTGGTGGCAGCGCCGACAGATTTGATCCATAGCTGAGAAAGTCTTTGATCATGCTGCT 1796
|||
Db 764 -----TyrAspAlaAspArgTyrThrGlnGlnGlnArgIleValTyrSerTyrLeuArg 781
|||
OY 1797 GGGACTTCTTATGTGTGTCGAGAAACGTGGGAGAGCCCTACCAAGGCTGATGATGTA 1856
|||
Db 782 TyrThrGlyThrProIleProAspAsnProAsnAspLysAsnAsnSerGlnLeuValVal 801
|||
OY 1857 GATTGT-----ACTTGCTGGGAGAGGCAAGCGACGATCACTTGACTTCAAGAAAT 1910
|||
Db 802 SerValAlaGlyThrValGlnGlyThrAsnGln-AspIleSerLeuLysPheGlnIle 821
|||
OY 1911 AGATGCAACGATCAGG 1926
|||
Db 821 eAspLeuThrSerArg 826
|||
```

Search completed: January 28, 2006, 02:42:08
Job time : 57.9006 secs


```

QY 648 AGCGTGTGCTGTTGAGGGGAGCAATCAAGACATTAGTCTTAATTTTGAATC 707
DB 21 ServAlaIaGIyThrValIuGIyThrAsnGlnAspIleSerLeuysPhepegIuile 40
QY 708 GATCTAAGATCAAGACCTGCTCATGAGGAAAGACAGACGCTTAAGTCCAAATCA 767
DB 41 AspleuThrSerArgProIahIsgIyLyThrGlnGlnIyLeuSerProIySser 60
QY 768 AAACCATTTGCTACTGATAGTGGCGCGAGTGTCAATAACTTGAGAAAGCTGACTTA 827
DB 61 LysProPheAlaThrAspSerGIyAlaMetSerHiIeLyLeuGluIyLysAlaAspleuLeu 80
QY 828 AAGGCTATTCAAGACAAATTTGATCGCTAACGTCACAGTACAGACGACTTGTGAGTC 887
DB 81 LysAlaIleGIInGlnIuLeuIleAlaAsnValHisSerAsnAspAspIyrPheGluVal 100
QY 888 ATTGATTTTGCAGGATCAACCATTAAGTCTGATGCAAGGCAAGGCTTACTTGTGAC 947
DB 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGIyLysValIyThrPheAlaAsp 120
QY 948 AAAGATGTTGCTGTAACCTTGGCGAGCCAACTGTCCAAGAATTTTGTAGCCGACAT 1007
DB 121 LysAspGIySerValIThrLeuProThrGlnProValGIInGluPheLeuSerGIyHis 140
QY 1008 GTGCGCGTTAGACCATATTAAGAAAAACCAATACAAAACCAAGCAATCTGTATGTC 1067
DB 141 ValArgValArgProGIyLysGIyLysProIleGIInAsnGlnAlaLysSerValAspVal 160
QY 1068 GAATATCTGTACAGTTTACCTTAAACCTGATGACGATTTTGACAGGCTCCAA 1127
DB 161 GIuIyThrValGIInPheThrProLeuAsnProAspAspPheArgProGIyLeuLys 180
QY 1128 GATATTAAGCTATTGAAAACTAGCTATCGGTGACACCATCAACATCTCAAGATTACTA 1187
DB 181 AspThrLysLeuLeuLysThrLeuAlaIleGIyAspThrIleThrSerGIInIyLeuLeu 200
QY 1188 GCTCAGACCAAGACATTTTAAACAAAACCAAGGCTATACGATTTATGACGTGAC 1247
DB 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGIyLysThrIleLysGIuAspArg 220
QY 1248 TCCCAATCGTACATGACATGACATGATTTCCGATGAGATTTTACATGATGACAGG 1307
DB 221 SerSerIleValIThrHisAspAsnAspIlePheAlaGIyThrIleLeuProAspGIInGlu 240
QY 1308 TTTACTTACCGTGTAAAAAATCGGGAACAAGCTTTATAGGATCAATAAAAATCGGTCTG 1367
DB 241 PheThrIyArgValIyAsnAsnArgGIInAlaIyArgIleAsnLysLysSerGIyLeu 260
QY 1368 AATGAGAAATTAACAACAACATGATCTGTGAGAAATATTACTCTTAAAAAGG 1427
DB 261 AsnGIuGIuIleAsnAsnThrAspLeuIleSerGIuLysIyLysValIleuLysLysGIy 280
QY 1428 GAAAGCGGTATGATCCCTTATGATGACATCACTGTAACCTGTCACCAATCGCT 1487
DB 281 GIuLysProIyAspProPheAspArgSerHisLeuLysLeuPheThrIleLysIyVal 300
QY 1488 GATGTCATACCAAGAAATTTGCTAAAAAGTGAAGAGCTTTAAACAGTACGCAACTAAC 1547
DB 301 AspValAspThrAsnGIuLeuLeuLysSerGIuGIInLeuThrAlaSerGIuAsnGln 320
QY 1548 TTAGACTTCAGAGATTTATAGACTCTCGTATTAAGGCTTAACTACTTACCAACTCTC 1607
DB 321 LeuAspPheArgAspLeuLysAspProArgAspLysAlaLysLeuLeuLysThrAsnLeu 340
QY 1608 GATGCTTTTGGTATTAAGACATTAACCTTAACCTGAAAAAGTGAAGATTAACAGATGAC 1667
DB 341 AspAlaAspGlyIleMetAspIyThrLeuThrGIyLysValGIuAspAsnHisAspAsp 360
QY 1668 ACCAACCGTATCATACCGTTATATAGGCAAGCGACCGAAGAGAGAGATGCTACTTAC 1727
DB 361 ThrAsnArgIleIleThrValIyMetGIyLysArgProGIuGIyGIuAsnAlaSerIy 380

```

```

QY 1728 CATTAGCT 1736
DB 381 HisLeuAla 383

RESULT 2
US-09-940-235-2
; Sequence 2, Application US/09940235
; Publication No. US2003005921A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kamnara
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL, CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/09/940,235
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Streptococcus equisimilis
US-09-940-235-2

Alignment Scores:
Pred. No.: 2,44e-172 Length: 414
Score: 1984.00 Matches: 383
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 53.25% Indels: 0
DB: 3 Gaps: 0

US-09-940-235-12 (1-2096) x US-09-940-235-2 (1-414)
QY 588 ATTGCTGACCTGAGTGGCTGTAGACCGTCACTGTCTCAACAACAGCAATTGGTGT 647
DB 1 IleAlaGIyProGIuThrPheLeuAspArgProSerValAsnAsnSerGIInLeuVal 20
QY 648 AGCGTGTGCTGTAAGTGGGAGGAGCAATCAAGACATTAGTCTTAATTTTGAATC 707
DB 21 ServAlaIaGIyThrValGIuGIyThrAsnGlnAspIleSerLeuysPhepegIuile 40
QY 708 GATCTAAGATCAAGACCTGCTCATGAGGAAAGACAGACGCTTAAGTCCAAATCA 767
DB 41 AspleuThrSerArgProIahIsgIyLyThrGlnGlnIyLeuSerProIySser 60
QY 768 AAACCATTTGCTACTGATAGTGGCGCGAGTGTCAATAACTTGAGAAAGCTGACTTA 827
DB 61 LysProPheAlaThrAspSerGIyAlaMetSerHiIeLyLeuGluIyLysAlaAspleuLeu 80
QY 828 AAGGCTATTCAAGACAAATTTGATCGCTAACGTCACAGTACAGACGACTTGTGAGTC 887
DB 81 LysAlaIleGIInGlnIuLeuIleAlaAsnValHisSerAsnAspAspIyrPheGluVal 100
QY 888 ATTGATTTTGCAGGATCAACCATTAAGTCTGATGCAAGGCAAGGCTTACTTGTGAC 947
DB 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGIyLysValIyThrPheAlaAsp 120
QY 948 AAAGATGTTGCTGTAACCTTGGCGAGCCAACTGTCCAAGAATTTTGTAGCCGACAT 1007
DB 121 LysAspGIySerValIThrLeuProThrGlnProValGIInGlnIuPheLeuSerGIyHis 140

```

```
QY 1008 GTGCGGCTTGACCATTAATAAGAAAAACAATACAAACCAAGCGAAATCTGGTAGTGTG 1067
DB 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
QY 1068 GAATATACGTAGAGTTTATCCCTTAACCCGTGATGACATTTTCAACACCGGTCTCAA 1127
DB 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGluLysLys 180
QY 1128 GATATACGCTATTGAAAAACATAGCTATGCTGACACCCATCATCATCTCAAGAAATTACTA 1167
DB 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 200
QY 1188 GCTCAAGCACAAACATTTTAAACAAAACCAACGAGCTATACGATTATGAAAGTGCAC 1247
DB 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGluTyrThrIleTyrGluArgAsp 220
QY 1248 TCCTCAATCGTCACTGATGACATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 1307
DB 221 SerSerIleValAlaThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
QY 1308 TTATCTTACCGCTGTAAAAATCCGGGAAACAAGCTTATGATCAATAAAAAATCGTCTG 1367
DB 241 PheThrTyrArgValAlaLysAsnArgGluGlnAlaIleTyrArgIleAsnLysLysSerGlyLeu 260
QY 1368 AATGAAGAAATAAACAACACTGACCTGATCTTGAGAAATATTAAGTCTTAAAAAAGGG 1427
DB 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValIleLysLysGly 280
QY 1428 GAAACACCGTATGATCCCTTGTATCCGACATCTTGAACCTGTTACACATCAATACGTT 1487
DB 281 GluLysProTyrLysProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
QY 1488 GATGTGATTAACAACAATTGCTTAAAAAGAGACAGCTCTTAAACAGTAAAGGAAACGTAAC 1547
DB 301 AspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320
QY 1548 TTAGACTTCAAGATTTATACGATCTCGATTAAGGCTAAACTCTCAACATCACTC 1607
DB 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
QY 1608 GATGCTTTTGGTATTATGACATTAACCTTAACTGAAAGTAGAGATTAATCAGATGAC 1667
DB 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
QY 1668 ACCAACCGTATCAATACCGTTTATAGGGACACGACCCGAAGAAGAGATGCTAGCTAC 1727
DB 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGluAsnAlaSerTyr 380
QY 1728 CATTAAGCT 1736
DB 381 HisLeuAla 383
RESULT 3
US-10-300-215-252
; Sequence 252, Application US/10300215
; Publication No. US20030153043A1
; GENERAL INFORMATION:
; APPLICANT: CARR, Francis Joseph
; APPLICANT: ADAIR, Fiona Suzanne
; APPLICANT: HAMILTON, Anita Anne
; APPLICANT: CARTER, Graham
; TITLE OF INVENTION: METHOD FOR THE PRODUCTION OF
; FILE OF INVENTION: NON-IMMUNOGENIC PROTEINS
; FILE REFERENCE: MER-104-Con.1
; CURRENT APPLICATION NUMBER: US/10/300,215
; PRIOR FILING DATE: 2002-11-20
; PRIOR APPLICATION NUMBER: US 09/438,136
; PRIOR FILING DATE: 1999-11-10
; PRIOR APPLICATION NUMBER: WO PCT/GB98/01473
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: GB 9710480.6
; PRIOR FILING DATE: 1997-05-21
; PRIOR APPLICATION NUMBER: GB 9716197.0
```

```

; PRIOR FILING DATE: 1997-07-31
; PRIOR APPLICATION NUMBER: GB 9725270.4
; PRIOR FILING DATE: 1997-11-28
; PRIOR APPLICATION NUMBER: GB 9807751.4
; PRIOR FILING DATE: 1998-04-14
; PRIOR APPLICATION NUMBER: US 60/067,235
; PRIOR FILING DATE: 1997-12-02
; NUMBER OF SEQ ID NOS: 254
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 252
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Streptococcus equisimilis
US-10-300-215-252

Alignment Scores:
Pred. No.: 5,866-170 Length: 414
Score: 1958.00 Matches: 378
Percent Similarity: 98.96% Conservative: 1
Best Local Similarity: 98.69% Mismatches: 4
Query Match: 52.55% Indels: 0
DB: Gaps: 0

US-09-940-235-12 (1-2096) x US-10-300-215-252 (1-414)
QY 588 ATTCGTGACCTGAGTGCTGCTAGACCGCTCATCTGTCACAAACAGCAATGCTGTT 647
DB 1 IleAlaGlyProGluThrLysLeuAspArgProSerValAlaAsnSerGlnLeuVal 20
QY 648 AGCGTTGCTGTACTGTGTGAGGGGACGAATCAAGCATTAAGTCTTAAATTTTGAATC 707
DB 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluLe 40
QY 708 GATTTAATCAATCAACACTGCTCATGAGAGAAAGACAGCAAGGCTTAAGTCCAAATCA 767
DB 41 AspLeuThrSerArgProAlaHisGlyLysThrGluGlnGlyLeuSerProLysSer 60
QY 768 AAACATTTGCTACTGATAGTGGGGCGCATGTCACATTAACCTTGAAGAAAGCTGACTACTA 827
DB 61 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu 80
QY 828 AAGCTATTCAAGAAACAATTGATGCTTAACCTCCACAGTAAACGACTTCTTGAAGGTC 887
DB 81 LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspTyrPheGluVal 100
QY 888 ATTGATTTGGACAGATGCAACCATTACTGATCGAAAACGCAAGGCTTACTTTCCTGAC 947
DB 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
QY 948 AAAAGATGTTGGTAACTTCCGACCCCAACCTGTCCAAAGATTTTGGTAAAGGAGACAT 1007
DB 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
QY 1008 GTGCGGCTTGACCATTAATAAGAAAAACAATACAAACCAAGGAAATCGTTGATGTG 1067
DB 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
QY 1068 GAATATACGTAGAGTTTATCCCTTAACCCGTGATGACATTTTCAACACCGGTCTCAA 1127
DB 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGluLysLys 180
QY 1128 GATATACGCTATTGAAAAACATAGCTATGCTGACACCCATCATCATCTCAAGAAATTACTA 1167
DB 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 200
QY 1188 GCTCAAGCACAAACATTTTAAACAAAACCAACGAGCTATACGATTATGAAAGTGCAC 1247
DB 201 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGluTyrThrIleTyrGluArgAsp 220
QY 1248 TCCTCAATCGTCACTGATGACATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 1307
DB 221 SerSerIleValAlaThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
```

Qy	1308	TTTACTTAACCGTTAAAAATCGGGAACAACCTTATAGGACATATAAAATCTGCTCG	1367
Db	241	PheMTrrYrhISvAllyLysaSmrGclGlnAlaTyrgIuIlleamLysLysSerGlyeu	260
Qy	1368	AATGAGAATAAACAACACTGACCTGATCTCTGGAAATATTACGTCCTTAAAAAGG	1427
Db	261	AenGluGluIlleamAnThrAspLeuIlleSerGluLysTyryValLeuLysLysGly	280
Qy	1428	GAAGAACCGTATGATCCCTTTGATCGAGTCACCTTGAACCTGTTACCATGAATACGTT	1487
Db	281	GluLysProTyrrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrrAl	300
Qy	1488	GATGTCGATCAACAAGAAATGTCTAAAAAGTAGACAGCTCTTAACAGCTAGGGAAGTAC	1547
Db	301	AspValAnThrAnGluLeuLeuLysSerGluGlnLeuThrIlaSerGluLysGln	320
Qy	1548	TTAGACTTCAGAGATTATACGATCTCTGTGATAAAGCTAAAGTAACTACTTCAACAATCTC	1607
Db	321	LeuAspPheAspLeuTyrrAspProArgAspLysAlaLysLeuLeuTyrrAspAnLeu	340
Qy	1608	GATGCTTTGGTATTTATGACCTAATACCTTAATCTGGAAGAGAGAGATTAACAAGATGAC	1667
Db	341	AspAlaIlePheGlyIleMetAspTyrrThrLeuThrGlyLysValGluAspAsnHisAspAsp	360
Qy	1668	ACCAACCGTATCATTAACCGTTATATAGGCAAGCAACCCGAAGAGAGAAATGTCTAGCTAC	1727
Db	361	ThrAsnArgIleIleThrValTyrrMetGlyLysArgProGluGluGluAsnAlaSerTyrr	380
Qy	1728	CATTTAGCT	1736
Db	381	HisLeuAla	383

RESULT 4

```

US-10-988-943-1
: Sequence 1, Application US/10988943
: Publication No. US20050176085A1
: GENERAL INFORMATION:
: APPLICANT: Center for Genetic Engineering and Biotechnology
: TITLE OF INVENTION: METHOD OF SELECTIVE PEPTIDE ISOLATION FOR THE
: TITLE OF INVENTION: IDENTIFICATION AND QUANTITATIVE ANALYSIS OF PROTEINS IN
: FILE OF INVENTION: COMPLEX MIXTURES.
: FILE REFERENCE: Proteomics CU2003-269
: CURRENT APPLICATION NUMBER: US/10/988,943
: CURRENT FILING DATE: 2004-11-15
: NUMBER OF SEQ ID NOS: 61
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 1
: LENGTH: 415
: TYPE: PRT
: ORGANISM: Streptococcus equisimilis
US-10-988-943-1

```

Alignment Scores:

Pred. No.:	5.86e-170	Length:	415
Score:	1958.00	Matches:	378
Percent Similarity:	98.96%	Conservative:	1
Best Local Similarity:	98.69%	Mismatches:	4
Query Match:	52.55%	Indels:	0
DB:	5	Gaps:	0

US-09-940-235-12 (1-2096) X US-10-988-943-1 (1-415)

QY	588	2	648	22	QY	708	42
TTTGCTGGACCGAGAGGCGTGCAGACCGTCATCTGTCAACAACGCCAATTGGTGT	TTTGCTGGACCGAGAGGCGTGCAGACCGTCATCTGTCAACAACGCCAATTGGTGT	TTTGCTGGACCGAGAGGCGTGCAGACCGTCATCTGTCAACAACGCCAATTGGTGT	TTTGCTGGACCGAGAGGCGTGCAGACCGTCATCTGTCAACAACGCCAATTGGTGT	TTTGCTGGACCGAGAGGCGTGCAGACCGTCATCTGTCAACAACGCCAATTGGTGT	TTTGCTGGACCGAGAGGCGTGCAGACCGTCATCTGTCAACAACGCCAATTGGTGT	TTTGCTGGACCGAGAGGCGTGCAGACCGTCATCTGTCAACAACGCCAATTGGTGT	TTTGCTGGACCGAGAGGCGTGCAGACCGTCATCTGTCAACAACGCCAATTGGTGT
64	64	21	70	41	76	61	61
Db	2	11	648	22	QY	708	42
TTTGCTGGACCGAGAGGCGTGCAGACCGTCATCTGTCAACAACGCCAATTGGTGT	TTTGCTGGACCGAGAGGCGTGCAGACCGTCATCTGTCAACAACGCCAATTGGTGT	TTTGCTGGACCGAGAGGCGTGCAGACCGTCATCTGTCAACAACGCCAATTGGTGT	TTTGCTGGACCGAGAGGCGTGCAGACCGTCATCTGTCAACAACGCCAATTGGTGT	TTTGCTGGACCGAGAGGCGTGCAGACCGTCATCTGTCAACAACGCCAATTGGTGT	TTTGCTGGACCGAGAGGCGTGCAGACCGTCATCTGTCAACAACGCCAATTGGTGT	TTTGCTGGACCGAGAGGCGTGCAGACCGTCATCTGTCAACAACGCCAATTGGTGT	TTTGCTGGACCGAGAGGCGTGCAGACCGTCATCTGTCAACAACGCCAATTGGTGT
64	21	70	41	76	61	61	61
Db	2	11	648	22	QY	708	42
TTTGCTGGACCGAGAGGCGTGCAGACCGTCATCTGTCAACAACGCCAATTGGTGT	TTTGCTGGACCGAGAGGCGTGCAGACCGTCATCTGTCAACAACGCCAATTGGTGT	TTTGCTGGACCGAGAGGCGTGCAGACCGTCATCTGTCAACAACGCCAATTGGTGT	TTTGCTGGACCGAGAGGCGTGCAGACCGTCATCTGTCAACAACGCCAATTGGTGT	TTTGCTGGACCGAGAGGCGTGCAGACCGTCATCTGTCAACAACGCCAATTGGTGT	TTTGCTGGACCGAGAGGCGTGCAGACCGTCATCTGTCAACAACGCCAATTGGTGT	TTTGCTGGACCGAGAGGCGTGCAGACCGTCATCTGTCAACAACGCCAATTGGTGT	TTTGCTGGACCGAGAGGCGTGCAGACCGTCATCTGTCAACAACGCCAATTGGTGT
64	21	70	41	76	61	61	61

QY	768	AAACATTTGGTACTGATAGTGGGCGGATGTCACATTAACCTTGGAAAGCTGACTTACTA	827
Db	62	LysProhepaIatnIspSerGclYAlaMeCProhIsIuSleuGIuVAlaIspIleu	81
QY	828	AAGGCTATTCAAGAACAAATTGATCGCTAACGTCCACAGTAACGACGACTTGTAGGTC	887
Db	82	LysIatIleGIInGIInIleuIleAlaasnValHisSerAasnAspPtyrPheGIuVal	101
QY	888	ATTGATTTTGGACGATGCAACCATTACTGATGAAACGGCAGGTCTACTTTCCTGAC	947
Db	102	IleAspPhealAsSerAlaThrIleThrAspArgasnGIuVAlaTyPheAlaAsp	121
QY	948	AAACATGGTTGGGTAAACCTTGGCCGACCCCAACCTGGCCCAAGAAATTTTCTAAGCGGACAT	1007
Db	122	LysAspGIleSerValThrIleuProThrGIInProValGIInGIuIleuIleuSerGIuHis	141
QY	1008	GTGGCGGCTTAGACCATATTAAGAAAAACCAATACAAACCAAGGAAATCTGTTGATGTC	1067
Db	142	ValArgValArgProTyArgIuGIuIuSerProIleGIInasnGIInAlaIuSerValAspVal	161
QY	1068	GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACATTTCAAGCCAGGTCTCAA	1127
Db	162	GIuTyThrValAlaInIleThrProIleuAsnProAspAspPheArgProGIuIleuIuS	181
QY	1128	GATACTAAGCATTAATGAAAAACACTAGCATCGGTGACACCATCACTCTCAAGAAATTACTA	1187
Db	182	AspThrIuSleuIeulysThrIleuAlaIleGIuIuAspThrIleThrSerGIInIleuIeu	201
QY	1188	GCTCAAGCAACAAGCATTTTAAACAAAAACCAACCAGGCTATACGATTATGAACGTGAC	1247
Db	202	AlaGIInAlaGIInSerIleIeulAsnIuSerThrIleSProGIuTyThrIleTyGIuIuArgsp	221
QY	1248	TCCCTCATTCGTCATCTGACAAATGACATTTCCGTACGATTTTATACCAATGATCAAGAG	1307
Db	222	SerSerIleValAlaThrIuSAspAsnAspIlePheArgThrIleIleuProIeAspGIInIu	241
QY	1308	TTTACTTAACGGTGTAAAAAATCGGGAAACAAGCTTATAGATCAATAAAAAATCTGCTGTG	1367
Db	242	PheThrIuTyThrIuSValIuSAsnArgIuGIInAlaTyGIuIleuIuIuSleuIuSleuIeul	261
QY	1368	AATGAAGAATTAACCAACACTGACCTGATCTCTGAGAATATTAACGTCTTAAAAAGGG	1427
Db	262	AsnGIuGIuIleAsnAsnThrAspIleuIleSerGIuIuSlyTyTyValIleuIuSleuGIu	281
QY	1428	GAAGAAGCCGATATGATCCCTTGAATCGGACGTCACTTGAACCTGTTCACCATCAATATACCTT	1487
Db	282	GIuIuSProTyArgAspProPheAspArgSerThrIleuIuSleuIuSleuPheThrIleTyTyVal	301
QY	1488	GATGTCAATACCAAGCAATTCCTTAAAAAGTAGAGGCTCTTAACAGCTAGCGAAGCTAAC	1547
Db	302	AspValAlaSerThrAsnGIuIleuIleuIuSleuSerGIuGIInIleuIuThrIleAsSerGIuIuArgsn	321
QY	1548	TTAGACTTCAGAGATTTATATACGATCTCTGTGATTAAGGCTAAACTACTTCAACAATCTC	1607
Db	322	LeuAspPheArgAspIleuTyArgAspProArgAspIuSlyAlaIuSleuIeulTyArgAsnIleu	341
QY	1608	GATCCTTTTGGTATTAATGACTATACCTTAACTGGAAGAAATGATGAGATTAACAGATGAC	1667
Db	342	AspAlaPheGIuIleIeulAspTyThrIleuThrGIuIuSValAlaGIuAspAsnHisAspAsp	361
QY	1668	ACCAACCGTATCATTAACCGTTTATATGCGGCAAGCGACCCGAAGAGAGAAATGCTAGTAC	1727
Db	362	ThrAsnArgIleIleIeulThrValIuIeMeGIuIuSArgProGIuGIuGIuAsnAlaSerTy	381
QY	1728	CATTAGCT 1736	
Db	382	HisIeuAla 384	

RESULT 5

US-10-300-215-253
; Sequence 253, Application US/10300215
; Publication No. US20030153043A1

TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic polypeptide
US-09-919-703-12

Alignment Scores:

Pred. No.:	3,276-167	Length:	413
Score:	1928.00	Matches:	377
Percent Similarity:	98.95%	Conservative:	1
Best Local Similarity:	98.69%	Mismatches:	2
Query Match:	51.74%	Indels:	2
DB:	3	Gaps:	2

US-09-940-235-12 (1-2096) x US-09-919-703-12 (1-413)

```

QY 588 ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAAATTGGTTGTT 647
Db 1 IleaIaGIyProGIuTtPleuLeuAspArgProSeValaAsnSeSerGIuLeuValaI 20
QY 648 AGCGTTGCGTACTGTTGAGGGGAGCAATCAAGACATTAAGCTTAATTTTGAATC 707
Db 21 SeValaIaGIyThrVal--GIyInraSngInaSPiLeSeuLeuysPhepGluIle 39
QY 708 GATCTAACATCAAGCACTGCTCATGAGAGAAAGACAGCAAGCTTAAGTCCAAATCA 767
Db 40 AspLeuThrSerArgProIaHisGIyGIyLysThrGIuGIuIndIyLeuSerProIySer 59
QY 768 AAACCAATTTGCTACTGATAGTGGCGGAGATGTCAATAACTTGAGAAAGCTGACTTA 827
Db 60 LysProPheIaThraSPeSerGIyAlaMetSerHisIeIySeuGIuIySaIaAspLeu 79
QY 828 AAGGCTATTCAAGAACAAATTCGCTACGTCACAGTACAGCAAGCACTTGAAGTTC 887
Db 80 LysaIaIleGIuSngInaIleuIleIaAsnValHisSeSaAsnAspArgPheGIuVal 99
QY 888 ATTGATTTTGAAGCGATCAACCACTTACTGATCGAAAGCGCAAGCTCTTGGCTGAC 947
Db 100 IleAspPheIaSeSaSPaIaThriIeThraSPaGAsnGIySaValIyPheIaAsp 119
QY 948 AAAGATGGTTCGGTAACTTGGCGAACCCAACTGTCCAAGAATTTTGGTAAAGCGACAT 1007
Db 120 LysAspGIySeValThrLeuProIhInProValGIuGIuPheIeLeuSeuSerGIyHis 139
QY 1008 GTGCCCGTTAAGCACTATTAAGAAAAACCAATCAAAACCAAGCGAAATGTGATGTG 1067
Db 140 ValaGyValaArgProIyLysGIuLysProIleGIuAsnGIuIaIaLysSeSaValaSPaVal 159
QY 1068 GAATATACTGTACAGTTTACTCCCTTAACCCCTGATGAGATTTTGACACGAGTCCAA 1127
Db 160 GIuIyThrValGIuPheThraSPeThraSPeAsnProAspAspPheArgProGIyLeuLys 179
QY 1128 GATATTAAGCTATTGAAAAACATAGCTATCGGTGACACCATCAATCTCAAGATTACTA 1187
Db 180 LeuThrIySleuLeuIyThrIeulIaIleGIyAspThriIeThSeSerGIuIuLeu 199
QY 1188 GCTCAAGCAACAAGCAATTTTAAACAACCAACCAAGCTTATGCAATTAAGAGTGAAC 1247
Db 200 AlaGIuIaGIuSeSaIeIeLeuAsnIySaAsnHisSPeGIyThrThriIeThrGIuIaGAsp 219
QY 1248 TCCCAATCGTCACTGATGACAAATGCAATTTCCGTAAGATTTTACCAATGATCAAGAG 1307
Db 220 SerSeSaIleValThrHisAspAsnSPiIePheArgThriIeLeuProMetArgGIuIu 239
QY 1308 TTTACTTACCGTGTAAAAATCGGGAACAAGCTTATAGATCAATAAAAATCTGGCTG 1367
Db 240 PheThrIyArgValIySaSnaRGIuGIuIaIyArgIleAsnIySlySeSaGIyLeu 259
QY 1368 AATGAAGAATTAACAACACTGACCTGATCTGTGAAGAAATAT--TACGCTCTTAAAAAA 1424
Db 260 AsnGIuGIuIleAsnAsnThraSPeIuIleSeSaIeGIuIyThrIySaValIeLeuLys 279
QY 1425 GGGGAAAAGCGTATGATCCCTTTGATGCCAGTCACTTGAAGACTGTTCACATCAATAC 1484

```

```

Db 280 GIyGIuLysProIyThraSPeProPheAspArgSeHisIeLeuLysIePheThriIeLysIyR 299
QY 1485 GTTATGTCATACCAACGAATTTGCTAAAGTAGACGCTCTTAACAGCTTAGCCAGCT 1544
Db 300 ValaSPaValaSPiThraSngInaIleuLeuLysSeSaGIuGIuIndIyThraISeSaGIuIuArg 319
QY 1545 AACTTAAGCTTCAGAGATTTTATGACATCTGCTGATTAAGGCTTAACCTCTACAAACAT 1604
Db 320 AsnLeuAspPheAspLeuThraSPeProIyArgAspIySaIaLysIeLeuIyThraSPe 339
QY 1605 CTCGATGCTTTGGTATTAATGACTATACTTAACCTGAGAAAGTAGAGATATACGAT 1664
Db 340 LeuAspIaPheGIyIleMetAspThrThriLeuThrGIyLysValGIuAspAsnHisAsp 359
QY 1665 GACACCAACCGTATCATACCTGTTATATGGGCAAGCAAGCAAGCAAGCAAGCAATCTGAC 1724
Db 360 AspThraSnaRGIyIleIeThraValIyMetGIySaArgProGIuGIuIaAsnIaSe 379
QY 1725 TACCAT 1730
Db 380 TyrHis 381

RESULT 7
US-10-474-792-658
; Sequence 658, Application US/10474792
; Publication No. US20040236072A1
; GENERAL INFORMATION:
; APPLICANT: Olmeda, Stephen
; APPLICANT: Zagursky, Robert
; APPLICANT: Nickbarg, Elliot
; APPLICANT: Winter, Louie
; TITLE OF INVENTION: SURFACE PROTEINS OF STREPTOCOCCUS PYOGENES
; FILE REFERENCE: AM 100399
; CURRENT APPLICATION NUMBER: US/10/474,792
; CURRENT FILING DATE: 2003-10-14
; NUMBER OF SEQ ID NOS: 674
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 658
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes
US-10-474-792-658

Alignment Scores:
Pred. No.: 6,226-155 Length: 440
Score: 1794.00 Matches: 341
Percent Similarity: 93.99% Conservative: 19
Best Local Similarity: 89.03% Mismatches: 23
Query Match: 48.15% Indels: 0
DB: 5 Gaps: 0

US-09-940-235-12 (1-2096) x US-10-474-792-658 (1-440)
QY 588 ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAAATTGGTTGTT 647
Db 27 IleaIaGIyYrGIyTtPleuProAspArgProIleAsnSeSerGIuLeuValaI 46
QY 648 AGCGTTGCGTACTGTTGAGGGGAGCAATCAAGACATTAAGCTTAATTTTGAATC 707
Db 47 SerMetIaIaGIyIleValGIuGIuIyThraSPeLysIySaValPheIleAsnPhePheGIuIle 66
QY 708 GATCTAACATCAAGCACTGCTCATGAGAGAAAGACAGCAAGCTTAAGTCCAAATCA 767
Db 67 AspLeuThrSerGIuProIaHisGIyGIyLysThrGIuGIuIndIyLeuSeuProIySer 86
QY 768 AAACCAATTTGCTACTGATAGTGGGAGATGTCAATAACTTGAGAAAGCTGACTTA 827
Db 87 LysProPheIaThraSPeAsnGIyAlaMetProHisIySleuGIuIySaIaAspLeu 106
QY 828 AAGGCTATTCAAGAACAAATTCGCTACGTCACAGTACAGCAAGCACTTGAAGTTC 887
Db 107 LysaIaIleGIuSngInaIleuIleIaAsnValHisSeSaAsnAspArgIyThraPheGIuVal 126

```

```

QY      888 ATGATTTTGGCAAGCGATCAACCAATTACTGATGAAACGGCAAGGTCTAATTGGCTGAC 947
      127 ILeaSPheAlaSerSPheAlaThrIleThraSPheArgAsnGlyValTyRheAlaAsp 146
QY      948 AAAAGATGGTCGGTAACCTTCCGACCCCAACCTGTCCAGAAATTTTGGCTAAGCGACAT 1007
      147 LysaSPGlySerValThreupProthrgInProValGlnGlnIubheuleuylsGlyHis 166
QY      1008 GTGCGCGTTAGACATATTAAGAAAAACAATAACAACCAACGAAATGTTGATGTCG 1067
      167 ValAlaGValArgProTyArgGlyulysProValGlnAsnGlnAlaLysSerValAspVal 186
QY      1068 GAATATCTGTACAGTTTATCCCTTAAACCTGTATGACGATTTTACAGACCGAGGTCTCAA 1127
      187 GluTyThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 206
QY      1128 GATACTAAGCTTTGAAAAACATAGCTATCGGTGACACCATGACATCTCAAGAAATTACTA 1187
      207 AspThrLysLeuLeuLysThreulAlaIleGlyAspThrIleThreSerGlnGlnLeu 226
QY      1188 GCTCAGACGACAAAGCATTTTAAACAACCAACCGGCTATCGATTTATTAAGACGTGAC 1247
      227 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyRThrIleTyGlnArgAsp 246
QY      1248 TCCCTCAATCGTCACTGATGACATGACATTTTCCGATACGATTTTACCAATGATCAAGAG 1307
      247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGln 266
QY      1308 TTTACTTACCGGTGTTAAAAATCGGGAACAAGCTTATAGATCAATAAAAATCGTGTCTG 1367
      267 PheThrTyRHisValLysAsnArgGlnGlnAlaTyGlnIleAsnProLysThrGlyIle 286
QY      1368 AATGAAGAAATAACAACAACCTGATCTGTGAATAATATTAGCTCTTAAAAAGG 1427
      287 LysGlyLysThrAsnAsnThrAspLeuValSerGlnLysTyRValLeuLysGlnGly 306
QY      1428 GAAAGCGGTATGATCCCTTTGATGCGAGTCACTTGAACCTGTTCAACAAATCGTT 1487
      307 GlnLysProTyArgProPheAspArgSerHisLeuLysLeuPheThrIleLysTyRVal 326
QY      1488 GATGTGATACCAACGAATTTGCTAAAAAGTAGACAGCTCTTAAACAGTACGGAACCTAAC 1547
      327 AspValAsnThrAsnGlnLeuLeuLysSerGlnGlnLeuThrIleAspGlnLysGln 346
QY      1548 TTAAGCTTCAGAGATTTATACATCTCTGTGATAGGCTTAACCTTCAACAAATCTC 1607
      347 LeuAspPheArgAspLeuTyArgProArgAspLysAlaLysLeuLeuTyRAsnAsnLeu 366
QY      1608 GATGCTTTTGTATATGACATATACCTTAACTGGAAGAGTAAAGATATCAACGATGAC 1667
      367 AspAlaPheAspIleMetAspTyRThrLeuThrGlyLysValGlnLysAsnHisAspLys 386
QY      1668 ACCAACCGTATACATACCTTTATATGGCAAGCGACCGCAAGAGAGAGATGCTAGCTAC 1727
      387 AsnAsnAlaGValValThrValTyRMetGlyArgArgProLysAlaLysGlySerTyR 406
QY      1728 CATTTAGCT 1736
      407 HisLeuAla 409

```

```

; PRIOR FILING DATE: 2002-05-24
; NUMBER OF SEQ ID NOS: 309
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 235
; LENGTH: 2355
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: sequence of fibrinectin
US-10-360-101-235

Alignment Scores:
Pred. No.: 7,45e-49 Length: 2355
Score: 638.50 Matches: 208
Percent Similarity: 36.69% Conservative: 69
Best Local Similarity: 27.55% Mismatches: 151
Query Match: 17.14% Indels: 328
DB: Gaps: 28

US-09-940-235-12 (1-2096) x US-10-360-101-235 (1-2355)

QY      5 GACCATTCATGTTGTCTGACAGTCCGACAGCTTTTGACAGACAGTCCCTTACAGT--- 61
      91 AspCysThrCysIleGlyAlaGlyArgGlyArgIleSerCysThrIleAlaAsnArgCys 110
QY      62 -----TCGTCGCGGTATCGGTGATTCATCTGCTAACCGTAAGGCAAC 106
      111 HisGlnGlyGlnGlnSerTyR-LysIleGlyAspThrTrp-----ArgArgPr 126
QY      107 CCGCAGCTTACGCGGCTCTCAACGACAGAGACAGATCATGCGCACCCGTGCGCAGA 166
      126 OHAsGlnThr----- 129
QY      167 CCAACGCTGCGCGAGATTCGATCCGCGAATTAATACGATCACTATAGGAGACCA 226
      129 ----- 129
QY      227 CAAGGTTCCCTAGAA---ATTAATTTGTTTAACCTTAAGAAGAGATATACATGG 283
      130 -GlyGlyTyRMetLeuGlnCysLeuGlyAsnGlyLysGlnTyRTrp---- 147
QY      284 TGCAAGACAAACAGATTGTACCATAGCTGAGAGAGTGTATGATCACTGCTGCGGACTT 343
      148 -----CysLysProIleAlaGlnLysCysPheAspHisAlaIleGlyThrS 163
QY      344 CTTATGTGTCGAGAAACGTGGA----- 368
      163 erTyRValValGlnGlnThrTrpGlnLysProTyRArgIleTyRMetMetValAspCysT 183
QY      369 -----GAAGGACGAGACGATCACTTGCATCTTGAAGAAATAGATGACAGATC 417
      183 hrCysLeuGlyGlnGlySerGlyArgIleThrCysThrIleSerArgAsnArgCysAsnAsp 203
QY      418 AGGACCAAGACATCTTAATAGAAATGGAGACACTGAGACCAAGAAATATATGAGAA 477
      203 LysPheThrArgHisSerTyRArgIleGlyAspThrTrpSerLysLysAspAsnArgGly 223
QY      478 ACCTGTCAGTCATCTGACAGGCAACGCGCAGAGAGAGTGAAGTGAAGAGCACA 537
      223 snLeuLeuGlnCysIleCysThrGlyAsnGlyArgGlyGlnTyRTrpLysCysGlnArgHisT 243
QY      538 CTTGTGTGACACCAACATGAGCGGATCTGAGCCCTTACCGATGTTTGTATTGCT--- 593
      243 hrSerValGlnThrThrSerSerGlySerGlyProPheThrAspValArgAlaAlaValT 263
QY      593 ----- 593
      263 yRglnProGlnProHisProGlnProProTyRArgHisCysValAlaThrAspSerGlyV 283
QY      594 -----GGACCTGAGTGGCTGTAGACCTGCACTGTGCAACAACAGCAAT 639
      283 alValTyRSerValGlyMetGlnThrPheu-----LysThrGlnGlyAsnLysGlnM 300

```

```

RESULT 8
US-10-360-101-235
; Sequence 235, Application US/10360101
; Publication No. US20040009550A1
; GENERAL INFORMATION:
; APPLICANT: Moll, Gert N.
; APPLICANT: Leenhouts, Cornelis J.
; TITLE OF INVENTION: Export and modification of (poly) peptide in the lantibiotic way
; FILE REFERENCE: 2183-5673
; CURRENT APPLICATION NUMBER: US/10/360,101
; CURRENT FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: EP 02077060.8

```

```

Qy 640 TGGTTGATGGCTTGGTCTGTAAGGAGGAGCAATCAAGATTAAGTCTTAATTT 699
Db      ::::: |||
Qy 300 etLeuCyThrCysLeuGly----- 306
Qy 700 TTGAATTCATCTAATCAATCAAGCTGCTCATGAGGAAAGAGCAAGGCTTAAGTC 759
Db      ::::: |||
Qy 307 -----AenGlyValSer- 310
Qy 760 CAATATCAAAACCATTTGCTACTGATAGTGGCGCATGTCAATAAATCTTGAAAGCTG 819
Db      ::::: |||
Qy 310 ----- 310
Qy 820 ACTTACTAAAGGCTATTCAGAGCAATGATGCTAAGCTCAAGTAAAGCAAGTACTACT 879
Db      ::::: |||
Qy 311 -----CysGlnGluThrAlaValThrlGlnThrTyrlGlyLeuSerAsnArg 326
Qy 880 TTGAG-----GTCAATGATTTTGCAGCGATGCAACCATTAAGTCAAGAAACGGCAAG 933
Db      ::::: |||
Qy 326 LysIleProCysValLeuProPheThr-----TyraGlnGlyArgT 339
Qy 934 TCTACTTTGCT-----GACAAAGATGGTTCGGTAACTTCCGAGCCCAAC 978
Db      ::::: |||
Qy 339 hrPheTySerCysThrThrlGlnGlyArgGlnAspGlyHisLeuTrpCysSerThrThrs 359
Qy 979 CTGTCAAGAAATTTTGTCTAAGCGAGCATGTGCGGTTAGACCATATAAAGAAACGAA 1038
Db      ::::: |||
Qy 359 etraNlyrGlu----- 362
Qy 1039 TACAAAACCAAGCAAAATCTGTGATGTGGAATATAGTGAAGTTTACTCCCTTAACC 1098
Db      ::::: |||
Qy 363 --GlnAspGlnLysTySerPheCysThrAspHisThrVal----- 375
Qy 1099 CTGATGACGATTTGACAGCAAGGTCTCAAGATTAAGTAAAGCAAGTAAAGTAAAGT 1158
Db      ::::: |||
Qy 375 ----- 375
Qy 1159 GTGACACATCATCATCTCAGAAATTAAGTCAAGCAAGCAAGCAATTTTAAACAAACC 1218
Db      ::::: |||
Qy 376 -----LeuValGlnThrGlnLys----- 381
Qy 1219 ACCGAGGCTATACGATTTATGACAGCTGACCTCTCAATGCTCATGACATGACATTT 1278
Db      ::::: |||
Qy 382 -----GlyAsnSerAsnGlyAlaLeuGlyHis----- 390
Qy 1279 TCCGTACGATTTTACCAATGATCAAGAGTTTACTTACCGTGTAAAAATCGGAGCAAG 1338
Db      ::::: |||
Qy 391 -----PheProPheLeuTyraAsnHis----- 398
Qy 1339 CTTATAGGATCATATAAAATCTGCTGATGAAAGAAATTAACAACAGTGAAGTGAATCT 1398
Db      ::::: |||
Qy 399 -----AsnTyThrAspCysThrS----- 405
Qy 1399 CTGAGAAATATTAAGTCTCTTAATAAAGGAGAAACCCGTATGATCCCTTGAATGCAATC 1458
Db      ::::: |||
Qy 405 etrGlnGly-----ArgAspAspA 411
Qy 1459 ACTTGAAGAA--CTGTTCACCATCAAAATACGTTGATGTGCATACCAAC-----G 1503
Db      ::::: |||
Qy 411 smetLysTrpCysGlyThrThrlGlnAsnTyraAspAlaAspGlnLysPheGlyPheCysP 431
Qy 1504 AATTGCTTAAAGAGAGAGAGCTCTTAACAGCTAAGGAGCAAGTAACTTAAGCTTCAAGAGAT 1563
Db      ::::: |||
Qy 431 rometAlaAlaHisGlnGluLysCysThrThraAsnGlu----- 443
Qy 1564 TATACGATCTCGTGATAGGCTAAAGCTAACAACATCTCGATGCTTTTGGTATTA 1623
Db      ::::: |||
Qy 444 -----GlyValM 446
Qy 1624 TGAATATACCTTAATCTGAAAGAGAGATTAATCAGATGACCAACCCGTATCATTA 1683
Db      ::::: |||
Qy 446 et--TyraArgIleGlyAspGlnTrpAspLysGlnHisAspMetGlyHisMetMetArgC 465
Qy 1684 CCGTTTATATGGGCAAGGAGCCCGAAGAGAGAAATGCTAGTCAATTTAGCTGGTGGTG 1743

```

```

Db      ::::: |||
Qy 465 yethrCysValGlyAsn-----GlyArgG 473
Qy 1744 GCCAGCGGCAACAGATTGA-----CCATAGCTGAAGAGTGTTCATCATGCTGCTG 1797
Db      ::::: |||
Qy 473 LysIleThrTrpThrCysIleAlaTyrserGlnLeuArgAspGlnCysIle-----ValAspA 491
Qy 1798 GGACTTCTATATGTGTGCGAGAAAGCGGAGAGGCGCTTCAAGCGCTGATGATGAGTAG 1857
Db      ::::: |||
Qy 491 epIleThrTyraAsnValaAsnAspThrPheHisLysArgHisGlnGlnGlyHisMetLeuA 511
Qy 1858 ATTGACTTGCCTGGAGAGAGGAGGAGCGAGCATCACTTGCACTTGAATAATAGATGA 1917
Db      ::::: |||
Qy 511 encYserThrCysPheGlnGlnGlyArgGlyArgTrpLysCysAspProValaAspGlnCysG 531
Qy 1918 ACCATGAGACACACAGACATCTTATGAATTTGAAGACACCTGAGACAGAAAGATATATC 1977
Db      ::::: |||
Qy 531 lnsAspserGlnThrGlnCysTyrcysTyrlYargGlylGlyGlnThrPhisCysGln- 550
Qy 1978 GAGGAACCTGCTCCAGTGCATCTGCACAGGCAACGCGGAGAGAGTGAAGTGTAGA 2037
Db      ::::: |||
Qy 551 --GlyValArgTyrlGlnCysTyrcysTyrlYargGlylGlyGlnThrPhisCysGln- 569
Qy 2038 GGCACACCTGTGTGACAGACACATCGAGCGGAGTCTGGCCCC 2078
Db      ::::: |||
Qy 570 -----ProLeuGlnThrTyrcysSerSerGlyPro 580

RESULT 9
US-09-961-403-1
; Sequence 1, Application US/09961403
; Publication No. US20030077589A1
; GENERAL INFORMATION:
; APPLICANT: HR-STUMPF, HOLGER
; APPLICANT: HAENDLER, BERNARD
; APPLICANT: KRAETZSCHMAR, JOERN
; APPLICANT: KREFT, BERTHOLD
; APPLICANT: WINTERHAGER, ELKE
; APPLICANT: REGIDOR, PEDRO
; APPLICANT: SCOTTI, SIMONE
; TITLE OF INVENTION: METHOD FOR IN VITRO DIAGNOSIS OF ENDOMETRIOSIS
; FILE REFERENCE: SCH-1789
; CURRENT APPLICATION NUMBER: US/09/961,403
; CURRENT FILING DATE: 2001-09-25
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2386
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-961-403-1

Alignment Scores:
Pred. No.: 7,49e-49 Length: 2386
Score: 638.50 Matches: 208
Percent Similarity: 36.69% Conservative: 69
Best Local Similarity: 27.55% Mismatches: 151
Query Match: 17.14% Indels: 328
Gaps: 28

US-09-940-235-12 (1-2096) x US-09-961-403-1 (1-2386)
Qy 5 GACCATTCATGTTGTTGCTCAGTGCAGACGTTTTCAGACAGAGTGCCTTACGT--- 61
Db      ::::: |||
Qy 122 AspCysThrCysIleGlyAlaGlyArgGlyArgIleSerCysThrIleAlaAsnArgCys 141
Qy 62 -----TCCGTCGGTATTCGTTGATTAATTCCTGAACCAATAGGCAAC 106
Db      ::::: |||
Qy 142 HisGlnGlyGlnInsertyr-LysIleGlyAspThrTrp-----ArgArgPr 157
Qy 107 CCGCCAGCTTAGCCGGGTCTTCAACAGACAGACAGATCATGCGACCCGTGGCCAGA 166
Db      ::::: |||
Qy 157 ohIsGlnThr----- 160

```


QY 167 CCCAAGCGTCCGAGATCTCGATCCGCGAATTAATACGACTATAGGAGACA 226
 Db 160 ----- 160
 QY 227 CAACGGTTCCCTTAGAA---ATAATTTGTTTAACTTTAAGAGATATACCATGG 283
 Db 161 -GlyGlyTyrMetLeuGlyCysValCysLeuGlyAsnGlyLysGlyGlyu-TyrThr----- 178
 QY 284 TGCAGACAAACAGATTGTACCTAGCTAGAGAGTTTGTATCATGCTGGGAGCTT 343
 Db 179 -----CysLysProIleAlaGlyLysCysPheAspHisAlaIleGlyThrS 194
 QY 344 CCTATGTGTGGAGAAACGTGGGA----- 368
 Db 194 eTyrValValGlyGlyuThrTrpGlyLysProTyrGlnGlyTyrMetMetValAspCysT 214
 QY 369 -----GAAGCAGCGAAGCATCATCTTGACCTTTAGAAATAGATGCAACGATC 417
 Db 214 hTyrLeuGlyGlyGlySerGlyArgIleThrCysThrSerArgAsnArgCysAsnArg 234
 QY 418 AGGACCAAGGAGATCTTATAGAAATTTGGAGACACCTGGACCAAGAAATATGAGAA 477
 Db 234 LAspThrArgThrSerTyrArgIleGlyAspThrTrpSerLysLysAspAsnArgGlyA 254
 QY 478 ACCGTCTCAGTCATCTGACAGCAAGCAGCGAGAGAGTGGAGAGTGTGAGAGACA 537
 Db 254 snLeuLeuGlnCysIleCysThrGlyAsnGlyArgGlyGlyuTrpLysCysGlnArgHisT 274
 QY 538 CCTGTGTGACAGACCAATCGAGCGGATCTGGCCCTTACACGATGTTCGTATTTGCT 593
 Db 274 hSerValGlnThrThrSerSerGlySerGlyProPheThrAspValArgAlaIleValT 294
 QY 593 ----- 593
 Db 294 TyrGlnProGlnProHisProGlnProProTyrGlyHisCysValThrAspSerGlyV 314
 QY 594 -----GGACCTGAGGTGGCTGTAGCCGTCCATCTGTCAACAAGCCCAT 639
 Db 314 aValTyrSerValGlyMetGlnTrpLeu-----LysThrGlnGlyAsnLysGlnM 331
 QY 640 TGGTGTGTAACGTTGCTGTACTGTGTGAGGGGAGCAATCAACATTAGTCTTAATTTT 699
 Db 331 eLeuCysThrCysLeuGly----- 337
 QY 700 TTGAATCGATTCAATCAACGACTGCTCATGAGAGAAAGACAGCAAGCTTAAGTC 759
 Db 338 -----AsnGlyValSer- 341
 QY 760 CAAATCAAAACCATTTGCTACTGATAGTGGCGGATGTCAATAACTTGAAGAAAGCTG 819
 Db 341 ----- 341
 QY 820 ACTTAAAGGCTATTCAAGAAATTTGATCGCTAAGCTCAACAGTAAAGCAGCATACT 879
 Db 342 -----CysGlnGlyuThrAlaValThrGlnThrTyrGlyGlyAsnSerArgS 357
 QY 880 TTGAG-----GTCAATTGATTTTTCAGCGCATCAACCATTAAGTATGAAACGCGAAG 933
 Db 357 LysGlnProCysValLeuProPheThr-----TyrAsnGlyArgT 370
 QY 934 TCTACTTTGCT-----GACAAAGATGTTGGTAACTTGGCCGACCAAC 978
 Db 370 hPheTyrSerCysThrThrGlnGlyArgGlnAspGlyHisLeuTrpCysSerThrThrS 390
 QY 979 CTGTCAAGAAATTTTGTAAAGCGGACATGTGGCGTTAGACCATATAAAGAAAAACAA 1038
 Db 390 eAsnTyrGlu----- 393
 QY 1039 TACAAACCAAGCAAGATCTGTGATGTGAATATATCTACAGTTACTCCCTTAACC 1098
 Db 394 --GlnAspGlnLysTyrSerPheCysThrAspHisThrVal----- 406
 QY 1099 CTGATGACGATTTCAAGACAGGTCTCAAAAGATACTAAGCTATTGAAAACTAGCTATCG 1158

Db 406 ----- 406
 QY 1159 GTGACACATGACATCTTAAGAAATTTACTAGCTCAAGCAACAAGCATTTTAAACAAAACC 1218
 Db 407 -----LeuValGlnThrGlnGly----- 412
 QY 1219 ACCAGGCTATACGATTTATGAGCGTACCTCTCAATGCTGATGCAATGACATTT 1278
 Db 413 -----GlyAsnSerAsnGlyAlaLeuCysHis----- 421
 QY 1279 TCCGTACGATTTTACCAATGATGACAGATTTACTTACTTCTGTTAAAAATCGGAAACAG 1338
 Db 422 -----PheProPheLeuTyrAsnHis----- 429
 QY 1339 CTATAGATCAATAAAAATCTGCTGAAATGAAAGAAATTAACAACATGACCTGATCT 1398
 Db 430 -----AsnTyrThrAspCysThrS 436
 QY 1399 CTGAGAAATATTACGCTCTTAAAAAGGGGAAAAAGCCGTATGATCCCTTGTATGCCAGTC 1458
 Db 436 eArgLysGly-----ArgArgAspA 442
 QY 1459 ACTTGAAG---CTGTTCAACATCAAAATAGCTTGTATGATGATACCAAC-----G 1503
 Db 442 snMetLysTrpCysGlyThrThrGlnAsnTyrArgAspLysAspGlnLysPheGlyPheCysP 462
 QY 1504 AATTGCTAAAAAGTGAAGAGCTCTTAAACAGTACGGAACGTAACCTTAGACTTCAGAGATT 1563
 Db 462 rMetAlaIleHisGlnGlyuTrpLecCysThrThrAsnGly----- 474
 QY 1564 TATACGATCTCGATGATGAGCTTAACCTATCTACAAACATCTGATGCTTTGGTATTA 1623
 Db 475 -----GlyValM 477
 QY 1624 TGAATTAACCTTAACGTGAAGAAAGTAAATACAGATGACACCAACGATCAATAA 1683
 Db 477 eC---TyrArgIleGlyAspGlnTrpAspLysGlnHisAspMetCysHisMetMetArgC 496
 QY 1684 CCGTTTATGAGGCAAGCAAGCCGAGAGAGAAATGCTAGCTACATTTAGCTGGTGGTG 1743
 Db 496 yThrCysValGlyAsn-----GlyArgS 504
 QY 1744 GCCAGGCGCAACAGATTGTA-----CCCATGCTGAGAGTGTTTGATCATGCTGCTG 1797
 Db 504 LysGluTrpThrCysIleAlaTyrSerGlnLeuArgAspGlnCysIle-----ValAspA 522
 QY 1798 GGACTTCCTATGTGTGGAGAAACGTGGAGAAAGCCCTACCAAGGCTGATGATGTAG 1857
 Db 522 ePileThrTyrAsnValAsnAspThrPheHisLysValGlnGlyuGlyHisMetLeuA 542
 QY 1858 ATTGTAATTTGCTGGAGAGAGCGAGCGGACGATCATCTTGCACTTCTGAATATGATGCA 1917
 Db 542 eAspThrThrPheGlyGlnGlyArgGlyArgGlyArgGlyArgGlyArgGlyArgGlyArgG 562
 QY 1918 ACGATCAGACACAAAGACATCTTAAATTTGAGACACCTGAGCAAGAAAGATTAATC 1977
 Db 562 LAspSerGluThrGlyThrPheThrGlnIleGlyAspSerTrpGlyLysTyrValHisS 581
 QY 1978 GAGGAAACCTGCTCAATGATCTGACAGGCAACGCGGAGAGAGAGTGGAGTGTAGA 2037
 Db 582 --GlyValArgTyrGlnCysTyrCysTyrGlyArgGlyIleGlyGlnTrpHisCysGln- 600
 QY 2038 GGCACACCTGTGTCAGACCAATCGAGCGGATGTGGGCC 2078
 Db 601 -----ProLeuGlnThrTyrProSerSerSerGlyPro 611
 Db 611 ----- 611
 RESULT 10
 US-10-788-792-206
 ; Sequence 206, Application US/10788792
 ; Publication No. US20040191819A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bayer Pharmaceuticals Corporation

APPLICANT: Eveleigh, Deepa
TITLE OF INVENTION: EXPRESSION PROFILES FOR BREAST CANCER AND METHODS OF USE
FILE REFERENCE: 5152
CURRENT APPLICATION NUMBER: US/10/788,792
CURRENT FILING DATE: 2004-02-27
PRIOR APPLICATION NUMBER: US 60/450,655
PRIOR FILING DATE: 2003-02-28
NUMBER OF SEQ ID NOS: 254
SOFTWARE: PatentIn version 3.2
SEQ ID NO 206
LENGTH: 2386
TYPE: PRT
ORGANISM: Homo sapiens
US-10-788-792-206

Alignment Scores:
Pred. No.: 7,49e-49 Length: 2386
Score: 638.50 Matches: 208
Percent Similarity: 36.69% Conservative: 69
Best Local Similarity: 27.55% Mismatches: 151
Query Match: 17.14% Indels: 328
DB: 4 Gaps: 28

US-09-940-235-12 (1-2096) x US-10-788-792-206 (1-2386)

QY 5 GACCATTCATGTTGTCAGGTCGACAGCGTTTGACAGACAGTCCGTTACGCT--- 61
DB 122 AspCysThrCysIleGlyAlaGlyArgGlyIleSerCysThrIleAlaSerAlaCys 141
QY 62 -----TCGTCGGGATCGGTGATTCGTTCTGTAACGTAAGGACCA 106
DB 142 HisGluGlyGlyGlnSerTyr-LysIleAlaAspIleTrp-----ArgAlaGly 157
QY 107 CGCAGCAGCTTACCGGGTCCGACGACGACGACGACATCGCACCGCGTGGCAGGA 166
DB 157 CHisGluTrp----- 160
QY 167 CCCAAGCGTCCGAGATCTCGATCCGGAATTAATAGACTCACTATAGGAGACCA 226
DB 160 ----- 160
QY 227 CAAGGTTTCCCTTAGAA---ATAATTTGTTTACTTTAAGAAGATATACATG 283
DB 161 -GlyGlyTyrMetLeuGlnCysValCysLeuGlnAsnGlyLysGln-TrpThr- 178
QY 284 TGCAAGCACAACAGATTGATCCCATAGTGAAGATTGATGCTGCTGGAGACT 343
DB 179 -----CysLysProIleAlaGlnLysCysPheAspHisAlaIleGlyThrS 194
QY 344 CCTATGTGTCGGAAGAACTGGGA----- 368
DB 194 eTyrValValIleGlyTrpGluLysProTyrGlnGlyTrpMetValAspCys 214
QY 369 -----GAAGCAGCGAGACGATGACTGCTTCTAGAAATATAGCAAGATC 417
DB 214 hrCysLeuGlnGlyIleGlySerGlyArgGlyLeuTrpCysThrSerArgAsnArgCysAsnAsp 234
QY 418 AGGACACAAGACATCTTATAGATTTGAGACACCTGAGCAAGAGATTAATCGAGAA 477
DB 234 InAspThrArgThrSerTyrArgIleGlyAspThrTrpSerTyrLysAspAsnArgGly 254
QY 478 AACTGCTCCAGTGCATCTGACAGGCAACGCGCAGAGAGAGTGAAGTGTAGAGGACA 537
DB 254 snLeuLeuGlnCysIleCysThrGlyAsnGlyArgGlyGlnTrpLysCysGlnArgHis 274
QY 538 CCTGTGTCAGGACCAATGAGAGCGATCTGGCCCTTCAACCATGTTCTGATTTGCT- 593
DB 274 hrSerValGlnThrThrSerSerGlySerGlyProPheThrAspValArgAlaAlaVal 294
QY 593 ----- 593
DB 294 yrgInProGlnProHisProGlnProProTyrGlnHisCysValThrAspSerGly 314

QY 594 -----GGACSTGAGTGGCTGTAGACCGTCCATCTGTCAACAAGCCAT 639
DB 314 alValTyrSerValGlyMetGlnTrpLeu-----LysThrGlnGlyAsnLysGln 331
QY 640 TGGTGTTCAGCGTTGCTGCTACTGTGTGAGGGGACGAATCAAGCATTAAGTCTTAATTT 699
DB 331 etLeuCysThrCysLeuGly----- 337
QY 700 TTGAATTCATCTTAACATCAACGACTGCTCATGAGGAGAAAGACAGCAAGCTTAAGTC 759
DB 338 -----AsnGlyValSer- 341
QY 760 CAAAATCAAAACCATTTGCTACTGATAGTGGCGGATGTCACTAAATCTGAGAAAGCTG 819
DB 341 ----- 341
QY 820 ACTTAACTAAGGCTATTCAAGAACATTTGATCGTAACTGACAGTAAAGACGACTACT 879
DB 342 -----CysGlnGlnThrAlaValThrGlnThrTyrGlyGlyAsnSerArg 357
QY 880 TTGAG-----GTCAATTGATTTTGCAAGCGATGCAACCATTACTGATCGAAACGGCAGG 933
DB 357 LysIleProCysValLeuProPheThr-----TyrAsnGlyArg 370
QY 934 TCTACTTTGCT-----GACAAAGATGTTGGGTAACTTCCGACCCGAC 978
DB 370 hrPheTyrSerCysThrThrGlnGlyArgGlnAspGlyHisLysLeuTrpCysSerThrS 390
QY 979 CTGTCCAAAGATTTTGTCTAAGCGGACATGTGCGGTTAGACCATATATAAGAAACCAA 1038
DB 390 exAsnTyrGlu----- 393
QY 1039 TACAAAACCAAGGAAATCTGTTGATGTGAATATATCTGACATTTATCCCTTAAAC 1098
DB 394 -GlnAspGlnLysTyrSerPheCysThrAspHisThrVal----- 406
QY 1099 CTGATGACGATTTCAAGCAGGTCTCAAAATACTAAGTATTTGAAAACTAGCTATCG 1158
DB 406 ----- 406
QY 1159 GTGACACCATCACATCTCAAGAAATTAAGTACGACGAAAGCATTTTAAACAAAC 1218
DB 407 -----LeuValGlnThrGlnGly----- 412
QY 1219 ACCGAGCTATAGATTTATGAACGTGACCTCAATCGTCACTGACATGACATGATTT 1278
DB 413 -----GlyAsnSerAsnGlyAlaLeuCysHis----- 421
QY 1279 TCCGTACGATTTTACCAATGATCAAGATTTTACTTACCGGTGTTAAAAATCGGAAACAG 1338
DB 422 -----PheProPheLysTyrAsnAsnHis----- 429
QY 1339 CTTATAGATCAATTAATAAATCTGCTGAATGAAGAAATTAACAACATGACCTGATCT 1398
DB 430 -----AsnTyrThrAspCysThrS 436
QY 1399 CTGAGAAATTAATACGTCCTTAAAAAGGGGAAAAAGCCGTATGATCCCTTGAATCGAGTC 1458
DB 436 eTyrGluGly-----ArgArgAsp 442
QY 1459 ACTTGAAGAA---CTGTTCACATCAAAATACCTTGAATGATCAAC-----G 1503
DB 442 snMetLysTrpCysGlyThrThrGlnAsnTyrAspAlaAspGlnLysPheGlyPheCysP 462
QY 1504 AATTGCTAAAAAGTGAAGCACTCTTAACAGTACGCAAGCTTAAGTGAAGTTCAGAGATT 1563
DB 462 rometAlaAlaHisGluGlnIleCysThrThrAsnGlu----- 474
QY 1564 TATACATCTCGTGAATTAAGCTTAACACTACTTCAACAAATCTCGATGCTTTGATTTA 1623
DB 475 -----GlyValM 477

```

OY 1624 TGAAGTAACTAACTTAATGGAAGAAAGATTAATCAAGATGAGACCAAGCGTATCATTA 1638
    ||| ||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Db 477 et---TyrArgIleelYarSerIntPrArybYsgInIiaArybSerelYhiIsmethcAtrc 496
OY 1664 CCGTTTATATGAGGCAAGGCAAGCCGGAAGAGAAATGATTAAGTACCAATTTAGCGTGGTGTG 1743
    ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
Db 496 ysthCyvAaIglYAn-----GIYArg 504
OY 1744 GCGAGGCGGCAACAGTTGTA-----CCATAGCTGAGAAAGTGTTTGATCATGCTGTG 1797
    ||| ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
    504 IYglutPrThCyvIlealYtrSerGIInleuAaYgAaRgIInCyvIle-----ValAaPa 522
OY 1798 GGAATTCCTATATGTGGTCCGAGAAACGTGGAGAAAGCCCTTACCAAGGCTGATGATGTAG 1857
    ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
Db 522 sPrlleThTyrAaInValIsmArThPrhEnIvYgAaRgIInGluGluuIYhiIsmethcAa 542
OY 1858 ATTGTACTCTGCGTGGGAGAAAGGCAAGCGGACGATCATCTTGCACTTTGAATAATGATGCA 1917
    ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
Db 542 snCyvsthCyvArhneIYglInIYArGrlYArGTrPrYCyvAaRPrvAlArRgInCyvG 562
OY 1918 ACGATTCAGGACAAAGACATCTATAGAAATTTGAGAACACCTGGAGCAAGAAAGATATC 1977
    ||| ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
Db 562 ImArbSerGIInThcIYhTrPheYTrGrlInIleGIYAaRSetrPrIuYsYTrValInIs- 581
OY 1978 GAGAAACCTGCTCCATGCAATCTTGCAAGGCAAGCGGCAAGAAAGTGAAGTGTGAGA 2037
    ||| ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
Db 582 --GIYValArGTrGInCyvYtrCyvYtrGrlYArGrlYIleGIYglutPrThsCyvGIn- 600
OY 2038 GGCACACCTCTGTGAGACCAACATTCAGGCGGATCTGGCCCC 2078
    ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
Db 601 -----ProIeuGIInThYtrProbSerSerGIYPro 611

```

```

Db      390  eTAsnTyrglu----- 393
QY      1039  TACAATAACCAAGCGAAATCTGTGATGTGGAATATATCTAGATTACTCCCTTAACC 1098
Db      394  --GlnAspDlnlySerPheCysThrAspHisThrVal----- 406
QY      1099  CTGATGAGATTTTCAGACCAAGGCTCAAGATATCTAAGCTATTGAAAACACTAGTATCG 1158
Db      406  ----- 406
QY      1159  GTGACACATCATCTCAGAAATTACTAGCTCAAGACCAAGCATTTTAAACAAAACC 1218
Db      407  -----LeuValGlnThrGlnGly----- 412
QY      1219  ACCCAGGCTATACGATTATGACGCTGCTCATCTGCTCATCTGACATGACATTT 1278
Db      413  -----GlyAsnSerAsnGlyAlaLeuCysHis----- 421
QY      1279  TCCGTAAGATTTTACCAATGATCAAGAGTTACTTACCGTGTAAATTCGGGAACAG 1338
Db      422  -----PheProPheLeuTyraAsnAsnHis----- 429
QY      1339  CTATAGATCATATAAAAATCTGCTGATGAAATGAAATTAACAACACTGACTGATCT 1398
Db      430  -----AsnTyrrThrAspCysThrS----- 436
QY      1399  CTGAGAAATATTACTCTCTTAAAAAGGGAAAAAGCCGTATGATCCCTTGTATCGCATC 1458
Db      436  eTglnGly-----ArgArgAsp 442
QY      1459  ACTGGAAT---CTGTTCAACATCAATACGTTGATGTGATACCAAC-----G 1503
Db      442  smMeLysTrpCysGlyThrThrGlnAsnTyraAspAlaAspGlnLysPheGlyPheCysP 462
QY      1504  AATGCTAAAAAGTGAAGAGCTCTTAACAGCTAGCGAAGTAAGTAAGTTCAGAGATT 1563
Db      462  rometAlaAlaHisGlnGlnLysCysThrThrAsnGlu----- 474
QY      1564  TATACGATCTCTGTGATTAAGCTAAACTCTTCAACAACATCTCGATGCTTTGTATTA 1623
Db      475  -----GlyValM----- 477
QY      1624  TGGACTATACCTTAACCTGAAAAAGTGAAGATATACAGATGACACCAACCGTATCATTA 1683
Db      477  et---TyraGrlLeuLysAspGlnTrpAspLysGlnHisAspMetGlnHisMetMetCysC 496
QY      1684  CCGTTATATAGGGCAAGCGAACCCGAAGAGAGATGTAGTACCATTTAGCTGGTGGTG 1743
Db      496  ysrThrCysValGlyAsn-----GlyArgG 504
QY      1744  GCCAAGCGCAACAGATTGTA-----CCCATAGCTGAAGATGTTTGAATCATGCTGCTG 1797
Db      504  LysGlnTrpThrCysLeuAlaTyrserGlnLeuArgAspGlnCysIle-----ValAsp 522
QY      1798  GGACTTCCTATGTGTGTGGAGAAACGTGGAGAACCCCTCAACAGGCTGGATGATGTGAG 1857
Db      522  spLleThrTyraAsnValAsnAspTrpPheHisLysArgHisGlnGlnGlyHisMetLeu 542
QY      1858  ATTGTAATCTGCTGGAGAGAGCGAGCGAGCATCATCTTCAAGAAATTAAGATGCA 1917
Db      542  snCysThrCysPheGlnGlnGlyArgGlyArgTrpLysCysAspProValAspGlnCysG 562
QY      1918  ACGATCAGACAGACAGACATCTTATAGAAATTGAGAGACACTGAGCAAGAAAGATTAAT 1977
Db      562  LlnAspserGlnThrGlnLysPheTyrglnLleGlyAspserTrpGlnLysTyrglnHis 581
QY      1978  GAGGAACCTCTCCAGTGCATCTGCACAGCGCAAGCGCCGAGAGAGATGGAAGTGTGAG 2037
Db      582  --GlyValArgTyrglnCysTyrglyArgGlyLleGlnGlnTrpHisCysGln- 600
QY      2038  GGCACACCTCTGTGACAGACATGAGCGGATCTGGGCC 2078
Db      601  -----ProLeuGlnThrTyrrProserSerSerGlyPro 611

```

```

RESULT 12
US-10-485-758-4
; Sequence 4, Application US/10485758
; Publication No. US20050129681A1
; GENERAL INFORMATION:
; APPLICANT: Varner, Judith A.
; TITLE OF INVENTION: Methods for Inhibiting Angiogenesis
; FILE REFERENCE: US00-07325
; CURRENT APPLICATION NUMBER: US/10/485,758
; PRIOR FILING DATE: 2004-02-03
; PRIOR APPLICATION NUMBER: 60/310,645
; PRIOR FILING DATE: 2001-08-06
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 2386
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-485-758-4

Alignment Scores:
Pred. No.: 7,49e-49 Length: 2386
Score: 638.50 Matches: 208
Percent Similarity: 36.69% Conservative: 69
Best Local Similarity: 27.55% Mismatches: 151
Query Match: 17.14% Indels: 328
Gaps: 28

US-09-940-235-12 (1-2096) x US-10-485-758-4 (1-2386)
QY      5  GACCAATCATGTTGTTGCTCAGTCCGACACGTTTTCGACAGACATGCTTACAGT--- 61
Db      122  AspCysThrCysLleGlyAlaGlyArgGlyArgLlserCysThrThrLeuAlaAsnArgCys 141
QY      62  -----TCGCTCGGCTATCGGTGATGTCATTCCTGCTAACCAATTAAGCAACC 106
Db      142  HisGlnGlyGlyGlnSerTyrrLysLleGlyAspPheTrp-----ArgArgPr 157
QY      107  CCGCAGCCTTAGCCGGGTCCTCAACGACAGACAGATCATGCGCACCCGTGGCCAGA 166
Db      157  ohisGlnThr----- 160
QY      167  CCCAAGCTGCCCCAGATCTCGATCCCGGAAATTAATACACTCACTATAGGAGACCA 226
Db      160  ----- 160
QY      227  CAACGGTTTCCCTTAGAA---ATAATTTGTTTACTTTAAGAGAGATATACATG 283
Db      161  -GlyGlyTyrrMetLeuGlnCysValCysLeuGlyAsnGlyLysGlyGlu---TrpThr--- 178
QY      284  TGCACACCAACAGATTGTATCCCATAGCTTGAAGAGTGTTTGATCATGCTGCTGGCACT 343
Db      179  -----CysLysProLleAlaGlnLysCysPheAspHisAlaLysLys 194
QY      344  CCTATGTGTGTGGAAGAACGTGGA----- 368
Db      194  eTyrrValValGlyGlnTrpGlnTrpGlnLysProTyrglnGlnTyrrMetMetValAspCys 214
QY      369  -----GAAGCAGCGGACCGATCATCTTGAAGAAATTAAGTGAACATC 417
Db      214  hrcysLeuGlnGlyGlnGlySerGlyArgLleThrCysThrSerArgAsnArgCysAsnAsp 234
QY      418  AGGACACAGACATCTTATAGAAATTGAGAGACACTGAGCAAGAAAGATTAATGAGAA 477
Db      234  LlnAspThrArgHisSerTyraGrlLeuLysAspThrTrpSerLysAspAsnArgGly 254
QY      478  ACCTGCTCAGTGCATCTGCACAGCGCAAGCGCCGAGAGAGATGGAAGTGAAGGACCA 537
Db      254  snLeuLeuGlnCysLleCysThrGlyAsnGlyArgGlyGlnTrpLysCysGlnArgHis 274
QY      538  CCTCTGTGACAGACCAATGAGCGGATCTGGCCCTTACCGAATGTCGATTTGCT--- 593

```

```

Db 274 hSeValGIlnThrThrSerSerglySerGIyProPheThrAspValAlaValAlaValT 294
Qy 593 -----
Db 294 yrgInProGIlnProHisProGIlnProProProTyGIlyHisCySValAlThAspSerglyV 314
Qy 594 -----
Db 314 aIValTySeValGIlyMetGIlnThrPleu -----LysThrInGIlnGlyAsnLysGIlnM 331
Qy 640 TGCGTTGTAACCGTGGCTGGTACTGTGTGAGGGAGCAATCAAGACATTAGCTTAAATTTT 699
Db 331 eLeuCySerThrCysLeuGly -----
Qy 700 TTGAATCGATCTAATCATCAGCAGCTGCTCATGAGAGAAAGACAGACAGCTTAAATGC 759
Db 338 -----
Qy 760 CAAAATCAAAACCATTTGGTACTGATAGTGGCGCGATGTCATAAACTTGAGAAAGCTG 819
Db 341 -----
Qy 820 ACTTACTAAAGCGTATTCAAGAACAATTGATCGCTAAAGCTCAACAGTAACAGCAGTACT 879
Db 342 -----
Qy 880 TTGAG-----GTCATTGATTTTGGCAAGCATCAACCATTAATCATGATCGAAACGGCAGG 933
Db 357 lYgluProCySValLeuProPheThr-----TyrAsnGlyAArgT 370
Qy 934 TCTACTTTGCT-----GACAAAGATGGTTCGCTAACTTCGCGACCCAGCAAC 978
Db 370 hPheTySerCySerThrThrInGIlnGlyAArgLysAspGlyHisLeuTyPyrSerThrThrS 390
Qy 979 CTGTCCAGAAATTTTGGCTAAGCGGACATGTGGCGCTTGAACCATTAAGAAACCA 1038
Db 390 eAsnTyrglu-----
Qy 1039 TACAAACCAAGCGAAATCTGTGATGTGAATATATCTGTACAGTTTACTCCTTAAAC 1098
Db 394 --GlnAspGlnLysTySerPheCySerThrAspHisThrVal-----
Qy 1099 CTGATGACGATTTCAAGACGAGTCTCAAGATATAGTATTAAGAAACCTAGATCG 1158
Db 406 -----
Qy 1159 GTGACACCATCATCTCAAGATTAATCTAGCTCAAGACACAAAGCATTTTAAACAAAC 1218
Db 407 -----
Qy 1219 ACCCAGGCTATACGATTTATGAACGTGACTCTCAATCGTCACTCATGACATGACATTT 1278
Db 413 -----
Qy 1279 TCCGTAAGATTTTACCAATGATCAAGATTTACTTACCGTGTAAATAACGGGAACAAG 1338
Db 422 -----
Qy 1339 CTTATAGATCAATAAAATCTGTCTGAATGAAGAAATTAACAACACTGACCTGATCT 1398
Db 430 -----
Qy 1399 CTGAGAAATATTATCGTCTTAAAGGGGAAAGCGGATGATCCCTTGTGATCGAGTC 1458
Db 436 eArgLysGly-----
Qy 1459 ACTTGAAG-----CTGTTACACATCAATACGTTGATGTGATCAACAC-----G 1503
Db 442 smMeLysTrpCySglYThrInGlnAsnTyrgLysAspAlaAspGlnLysPheGlyPheCyS 462
Qy 1504 AATTCTTAAAGAGTGAAGCTCTTAAACAGTACGTAAGCAAGTAACTTAAGACTTGAAGATT 1563
Db 462 rMeAlaAlaHisGlnGlnLysCysThrThrAsnGlu-----

```

```

Qy 1564 TATACATCTCGTATGAAGGCTAAAGCTAACTCTACAAACATCTGATCTTTGGTATTA 1623
Db 475 -----
Qy 1624 TGACATATACCTTAATCTGAAAAGTGAAGATTAATCAAGATGACACCAACCGTATCATTA 1683
Db 477 eC--TyrArgLleGlyAspGlnThrAspLysGlnHisAspMeGlyHisMetMetAlaGc 496
Qy 1684 CCGTTATATGGGCAAGCGACCCGAAAGAGAAATGCTAGCTACATTAGCTGTGTGTG 1743
Db 496 yThrCySValGlyAsn-----
Qy 1744 GCCAGGCGCAACATATTGTA-----CCATAGCTGAGAGTGTGTTGATCATGCTGTG 1797
Db 504 lYgluTrpThrCySleAlaTySerGlnLeuAlaAspGlnCySle-----ValAspA 522
Qy 1798 GGACTTCTATATGTTGGTGGAGAAACGTGGAGAAAGCCCTTACCAAGGCTGATGATGATG 1857
Db 522 ePleThrTyrgLysValAsnAspThrPheHisLysValArgHisGlnGlnGlyHisMetLeuA 542
Qy 1858 ATTGTACTTGGCTGGAGAAAGCGAGCGCATCATCTTGAATTAAGAAATAGATCA 1917
Db 542 snCySerThrCysPheGlyGlnGlyAArgGlyAArgTrpLysCyAspProValAspGlnCySg 562
Qy 1918 ACATCAGACACACAGACATCTTATGAATTTGAGACACCTGAGACAAAGATATATC 1977
Db 562 lAspSerGluThrGlyThrPheTyrgLleGlyAspSerTrpGlnLysTyrgValHisS 581
Qy 1978 GAGAAACCTGCTCTCACTGATCTGACACAGGCAACGCGCGAGAGAGTGAAGTGTGA 2037
Db 582 --GlyValAlaGlyTyrgLincySerTyrgLysAArgGlyLleGlyGlnTrpHisCySgIn- 600
Qy 2038 GGCACACCTGTGTCAGACACCATGACGCGATCTGGCCCC 2078
Db 601 -----

```

RESULT 13

```

US-10-485-758-9
; Sequence 9, Application US/10485758
; Publication No. US20050129681A1
; GENERAL INFORMATION:
; APPLICANT: Varmer, Judith A.
; TITLE OF INVENTION: Methods for Inhibiting Angiogenesis
; FILE REFERENCE: USCD-07325
; CURRENT APPLICATION NUMBER: US/10/485,758
; CURRENT FILING DATE: 2004-02-03
; PRIOR APPLICATION NUMBER: 60/310,645
; PRIOR FILING DATE: 2001-08-06
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 2386
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-485-758-9

```

Alignment Scores:

```

Pred. No.: 7,49e-49 Length: 2386
Score: 636.50 Matches: 208
Percent Similarity: 36.69% Conservative: 69
Best Local Similarity: 27.55% Mismatches: 151
Query Match: 17.14% Indels: 328
DB: 5 Gaps: 28

```

US-09-940-235-12 (1-2096) x US-10-485-758-9 (1-2386)

```

Qy 5 GACATTCATGTTGTTGCTGAGTGGAGACGTTTTCAGACAGAGTGGCTTCAAGT--- 61
Db 122 AspCySerThrCysLleGlyAlaGlyAArgGlyAArgLysCysThrLleAlaAsnArgCys 141
Qy 62 -----TGGCTCGGCTATCGGTGATTCATTTCTGCTTAACAGTAAGGCAAC 106
Db -----

```

Db 142 HisGluGlyGlyHisSerTyr-LysIleGlyAspThrTyr-----ArgArgPr 157
 Qy 107 CGCCAGCCTTACCGGGTCTCTCAACGACAGACGATCATGGCACCCTGGCCAGCA 166
 Db 157 OHISGLUTHR----- 160
 Qy 167 CCCAACGCTGCCGAGATCTCGATCCCGGAAATTAAACGACTCTATAGGGAGACCA 226
 Db 160 ----- 160
 Qy 227 CAACGTTTCCTCTTGA--ATAATTTTGTAACTTTAAGAGAGATATACATGG 283
 Db 161 -GlyGlyTyrMetLeuGlyCysValCysLeuGlyAsnGlyGly-u-TyrThr- 178
 Qy 284 TGCAGACACACAGATTGATACCATAGTACGAGAAAGTTTGAATGATGCTGGGACT 343
 Db 179 -----CysIysProIleAlaGluIuScysPheAspHisAlaIleGlyThrs 194
 Qy 344 CCTATGTGTGTCGAGAAACGTGGG----- 368
 Db 194 eTyrValValGlyGluThrTrpGluLysProTyrGlnGlyTyrMetMetValAspCysT 214
 Qy 369 -----GAGGACACGCGACGATCATCTTGCATCTTGAATAATAGTCAACGATC 417
 Db 214 hrCysLeuGlyGlyGluGlySerGlyArgIleThrCysThrSerArgAsnArgCysAsnAspG 234
 Qy 418 AGGACACAGACATCTTATAGTAATTTGAGACACCTGGAGCAAGAGAGATATACGAGAA 477
 Db 234 LAspThrArgThrSerTyrArgIleGlyAspThrTrpSerLysAspAsnArgIYA 254
 Qy 478 ACCCTCTCAGTGCATCTGACAGAGCAAGCGCCGAGAGAGATGGAAGTGGAGAGGACA 537
 Db 254 snLeuLeuGlnCysIleCysThrGlyAsnGlyArgGlyGluTrpLysGlyIuArgHisT 274
 Qy 538 CCTGTGTGACACACATCGACGCGAGTGGCCCTTACCCGATGTTCTGATTTGCT- 593
 Db 274 hrSerValGlnThrThrSerSerGlySerGlyProPheThrAspValArgAlaIleValT 294
 Qy 593 ----- 593
 Db 294 YrGlnProGlnProHisProGlnProProTyrGlyHisCysValThrAspSerGlyV 314
 Qy 594 -----GACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAT 639
 Db 314 alValTyrSerValGlyMetGlnTrpLeu-----LysThrGlnGlyAsnLysGlnM 331
 Qy 640 TGGTTGTAGCGCTGTGCTGCTGTTGAGGGAGCGAATCAACATTAGTCTTAAATTTT 699
 Db 331 etLeuCysThrCysLeuGly----- 337
 Qy 700 TTGAATGCATCTAACATACGACCTGCTCATGAGAGAGAGAGAGAGAGAGAGAGAGAG 759
 Db 338 -----AsnGlyValSer- 341
 Qy 760 CAATCAAAACCATTTTGCTACTGATGATGAGCGCGATGTCACATAACTTGAAGAAAGCTG 819
 Db 341 ----- 341
 Qy 820 ACTTAAAGGCTATTCAAGAACATTTGCTTAAAGTCAACGTCACAGTAAACGACGACTACT 879
 Db 342 -----CysGlnGluThrAlaValThrGlnThrTyrGlyGlyAsnSerAsnG 357
 Qy 880 TTGAG-----GTCAATGATTTTTCAGAGGAGTCAACCATTTACTGATCGAAACGCGCAGG 933
 Db 357 LysGlnProCysValLeuProPheThr-----TyrAsnGlyArgT 370
 Qy 934 TCTACTTGTCT-----GACAAAGATGGTTGCGTAACCTTGGCCGACCCAC 978
 Db 370 hrPheTyrSerCysThrThrGluGlyArgGlnAspGlyHisIleuTrpCysSerThrThrs 390
 Qy 979 CTGTCCAGAAATTTTGTCTAGCGGACATGTGCGCTTAGACCATATTAAGAAAAACAA 1038
 Db 390 eTAsnTyrGlu----- 393

Qy 1039 TACAAACCAAGCGAAATCTGTGATGGAATATATCTGACATTTACTCCCTTAAAC 1098
 Db 394 --GlnAspGlnLysTyr-SerPheCysThrAspHisThrVal----- 406
 Qy 1099 CTGATGACGATTTTACACACAGGTCTCAAAAGATCTAATGATTAAGAAAACACTAGCTATCG 1158
 Db 406 ----- 406
 Qy 1159 GTGACACCATTCATCTCAGAAATTTACTAGTCAAGACCAAAAGCATTTTAAACAAAAC 1218
 Db 407 -----LeuValGlnThrGlnGly----- 412
 Qy 1219 ACCGAGCTATAGATTTTATGACGTGACTCTCAATGCTACCTGATGACATGACATTT 1278
 Db 413 -----GlyAsnSerAsnGlyAlaLeuCysHis----- 421
 Qy 1279 TCCGTAGCATTTTACCAATGATCAAGATTTACTTACCGTGTATTAATAATCGGAAACAG 1338
 Db 422 -----PheProPheLeuTyrAsnAsnHis----- 429
 Qy 1339 CTATATAGATCATTAATAAATCTGCTGTAATGAAGAAATTAACAACACTGACTGATCT 1398
 Db 430 -----AsnTyrThrAspCysThrs----- 436
 Qy 1399 CTGAGAAATATATAGTCTTAAATAAAGGGGAAAAAGCGGTATGATCCCTTGTGACGAGTC 1458
 Db 436 eTyrGluGly-----ArgArgAsp 442
 Qy 1459 ACTTGAAA--CTGTGACCATTCATTAATCGTTGATGTCATACCAAC-----G 1503
 Db 442 snMetLysTrpCysGlyThrGlnIleAsnTyrAspAlaAspGlnLysPheGlyPheCysP 462
 Qy 1504 AATGCTTAAAGTAGACGAGCTTTAACAGCTTAGCGAAGCGTAACTTGAATCCTTACAGAGATT 1563
 Db 462 rMetAlaAlaIleGluGluIleCysThrThrAsnGlu----- 474
 Qy 1564 TATATGATCCCTCGATTAAGGCTTAACCTACTCTACACAACTGATGCTTTGGTATTA 1623
 Db 475 -----GlyValM 477
 Qy 1624 TGCATATACCTTAACTGAAAAGTAGAGATTAATACGATGACACCAACCGTATCATTA 1683
 Db 477 et---TyrArgIleGlyAspGlnTrpAspLysGlnHisAspMetClyHisMetMetAlaG 496
 Qy 1684 CCGTTATATAGGCGAACGACCCGAGAGAGAAATGCTAGCTTACATTTAGCTGGGTG 1743
 Db 496 YsThrCysValGlyAsn-----GlyArgG 504
 Qy 1744 GCCAGGCGCAACAGATTGTA-----CCCATAGCTGAGAAAGTGTTTGATCATGCTGCTG 1797
 Db 504 LysIuTrpThrCysIleAlaTyrSerGlnLeuAspGlnCysIle-----ValAsp 522
 Qy 1798 GAACTTCTATGTGTGCGAGAAAGCTGGAGAAAGCCCTTACCAAGGCTGATGATGTAG 1857
 Db 522 spIleThrTyrAsnValAsnAspThrPheHisIleValGlnGluGlyHisMetLeuAla 542
 Qy 542 snCysThrCysPheGlyGlnGlyArgGlyArgTrpLysCysAspProValAspGlnCysG 562
 Db 1918 ACATGACGACACACAGACATCTTATGAATTTGAGACACCTGAGACAAAGAGATATC 1977
 Qy 562 LAspSerGluThrGlyThrPheTyrGlnIleGlyAspSerTrpGluLysTyrValHis- 581
 Db 1978 GAGGAAACCTGCTCAGTGCATCTGACAGGCAACGCGGAGAGAGAGAGAGAGAGAGAGAG 2037
 Db 582 --GlyValAlaGlyTyrGlnCysTyrCysTyrGlyArgGlyIleIleGluTrpHisCysGln- 600
 Qy 2038 GGCACACCTTGTGACAGACCAATCGACCGGATGGCCCC 2078
 Db 601 -----ProLeuGlnThrTyrProSerSerSerGlyPro 611

```

RESULT 14
US-10-236-392-2
; Sequence 2, Application US/10236392
; Publication No. US20040067490A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David M
; APPLICANT: Boldog, Ferenc L
; APPLICANT: Burgess, Catherine, E
; APPLICANT: Casman, Stacie J
; APPLICANT: Catberton, Elina
; APPLICANT: Chapoval, Andrei
; APPLICANT: Crabtree, Julie
; APPLICANT: Edinger, Shlomit, R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gerlach, Valerie
; APPLICANT: Gorman, Linda
; APPLICANT: Grosse, William M
; APPLICANT: Gusev, Vladimir
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Larocheille, William J
; APPLICANT: Li, Li
; APPLICANT: MacDougall, John R
; APPLICANT: Malpankar, Uriel M
; APPLICANT: Miller, Charles E
; APPLICANT: Millet, Isabelle
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Patturajan, Weera
; APPLICANT: Pena, Carol A
; APPLICANT: Peyman, John A
; APPLICANT: Rastelli, Luca
; APPLICANT: Reiger, Daniel K
; APPLICANT: Rottemberg, Mark E
; APPLICANT: Shenoy, Suresh
; APPLICANT: Shimkets, Richard A
; APPLICANT: Smithson, Glenda
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-442A
; CURRENT APPLICATION NUMBER: US/10/236.392
; CURRENT FILING DATE: 2002-09-06
; PRIOR APPLICATION NUMBER: US09/540,763
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: US60/390,155
; PRIOR FILING DATE: 2002-06-19
; PRIOR APPLICATION NUMBER: US09/635,949
; PRIOR FILING DATE: 2000-08-10
; PRIOR APPLICATION NUMBER: US60/318,765
; PRIOR FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: US60/357,303
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US60/367,753
; PRIOR FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER: US60/369,479
; PRIOR FILING DATE: 2002-04-02
; PRIOR APPLICATION NUMBER: US09/659,634
; PRIOR FILING DATE: 2000-09-12
; PRIOR APPLICATION NUMBER: US60/318,120
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US60/318,130
; PRIOR FILING DATE: 2001-09-07
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 794
; SOFTWARE: Custom
; SEQ ID NO 2
; LENGTH: 2320
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-236-392-2

Alignment Scores:
Pred. No.: 1,266-48 Length: 2320
Score: 636.00 Matches: 205
Percent Similarity: 36.56% Conservative: 71
Best Local Similarity: 27.15% Mismatches: 154

```

Query Match:	17.07%	Indels:	326
DB:	4	Gaps:	23
US-09-940-235-12 (1-2096) x US-10-236-392-2 (1-2320)			
QY 5 GACGATTTATGTTGTTGCTCAGGTGGCAGACGTTTGGACAGACAGACGTCCTACAGT---			61
DB 87 AWPCTHNCYAILLEGLVALAGLYAGIILSEFCYTHRIILEALAAHAYGYS			106
QY 62 -----TCGCTCGGATTCGGTGAATTCATTCCTGTAACAGTAAGGCAACC			106
DB 107 HIsGLUGLYGLHsERTYr-LysELLEGLYsHPHTRP-----ArgHsGr			122
QY 107 CCGCAGGCTTACCGGGTCTTCAAGCAGAGACAGATCATGCGACCCGTGGCAGGA			166
DB 122 OHIsGLuThr-----			125
QY 167 CCGAAGCGTCCGAGATCTCATCCCGGAAATTAATACACTCATATAGGAGACCA			226
DB 125 -----			125
QY 227 CAACGGTTTCCTCTAGAA--ATATATTTGTTTAATTAGAGAGAGATATACATGG			283
DB 126 -GLYGLYrYrmetLeuGLuCYsVALCYsLeuGLYsHsGLYGLYGLu-TrpThr---			143
QY 284 TGCAGACACACAGATTTGATCCCATATGCTGAGAGAGTTTGTATCATCTGCTGGACTT			343
DB 144 -----CYsLYsPrOILleAGLULYsCYsPHeAsPHIsAlleAGLYThrs			159
QY 344 CCTATGTGTGGAGAAACGTGGAA-----			368
DB 159 ERTYrVALVALGLYGLuThrTrpGLYsPrOYrGLYGLuLYrTrpMetMetVALsPrCYsT			179
QY 369 -----GAAGGACGCGAGCGCATCTTGCACTTGAATAATAGATGCAACGATC			417
DB 179 hrcYsLeuGLYGLuGLYsERTYrYAGIILEThrcYsThrsErtAgaHAYGYSAsHsPc			199
QY 418 AGGACACAAAGACATCTTATAGAAATTGAGAACCTGAGCAAGAAAGATATCGAGAA			477
DB 199 LhAsPThrArghrThrsERTYrAGIILEGLYAsPThrTrpERTYrLYsAsPAsHsAGLYA			219
QY 418 AACTGCTCCAGTGCATCTTGCACAGGCAAGCGCCGAGAGAGATGAGAGGACCA			537
DB 219 snLeuLeuGLuCYsIIECYsThrGLYsHsGLYGLYGLuTrpLYsCYsGLuATGHIsT			239
QY 538 CCTGTGTGACAGCAATGAGCGGATCTGGCCCTTACCCGATGTTCTGATTTGCT---			593
DB 239 hTrsVALGLuThrThrsERTYrGLYsERTYrPrOphEThrAsPVALsAlleVALIT			259
QY 593 -----			593
DB 259 YrGLhPrOGLhPrOHIsPrOGLhPrOPrOTrOTrGLYHIsCYsVALThrAsPSErGLYV			279
QY 594 -----GACCTGAGTGGCTGCTGACCGCTCCATCTGTGCACACAGCCAAAT			639
DB 279 alVALYrSErVALGLYMetGLuTrpLeu-----LYsThrGLuGLYsHsGLuIM			296
QY 640 TGGTTGTAGCGCTTCTGCTGATCTGTGAGGGGACGAGATCAAGACATTAAGCTTAATTTT			699
DB 296 eTLeuCYsThrcYsLeuGLYsHsGLYsErcYsGLuGLuThrAlaVALThrGLuThrT			316
QY 700 TTGAAATGCATTAATCATCAGCACT--			744
DB 316 YrGLYGLYAsHsLeuHsHsGLYGLuPrOCYsVALLeuPrOphEThrTrYrAsHsGLYAGThPr			336
QY 745 AGCAAGGCTTAAGTCCAAATCAAAACCATTTGTCTAGTATGTCGGCGAGTCCATTA			804
DB 336 hErYrSErCYsThrThrGLYAGGLu-----			345
QY 805 AACTTGAAGAAAGCTGACTTCTAAAGGCTATTCAGAAACAATTTGATCGCTAACCTCCAC			864
DB 346 -----AsPGLYHIsLeuTrpCYsSErThrThrs			355

```

QY      865 GTAAGCAGCACTACTTTGAGTCATTGATTTTGCAGCGATGCAACCATTAAGTCAATGCA  924
      |||||  ::  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      355 eTAhntYrGluInlmerGlnlyrSerPheCyserThrAspHisThrVal-LeuValGln  374
QY      925 ACGGCAAGGTCTACTTTGCTGCAAAAGATGGTTCGGTAACTTTGCCGAGCCCAACTGGCC  984
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      375 Thr-GlnGly-----GlyAsnSerAsnGlyValLeuCyS--  385
QY      985 AAGATTTTGTGTAAGCGGACATGTCGGCTTAGACCATTAAGAAAAACCAATACAA  1044
      385 -----
QY      1045 ACCAAGCGAAATCTGTGATGATGATATATAGTACAGTTACTCCCTTAAACCTGATG  1104
      385 -----
Db      385 -----
QY      1105 ACGATTTGACAGCAGGTCTCAAGATATAGCTATTGAAACATAGCTATGCGTGACA  1164
      385 -----
Db      385 -----
QY      1165 CCATCATCTCTAAGATTAAGTCTCAAGCACAAGACATTTTAAACAAACCAACCCAG  1224
      385 -----
Db      385 -----
QY      1225 GCTATACGATTTATGAAAGTGACTCCTCAATGCTGATGACATGACATTTTCCGTA  1284
      386 -----
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      386 -----HisPheProPh  389
QY      1285 CGATTTTACCAATGATCAAGAGTTTACTACCGTGTAAATCCGGAACAAGCTTATA  1344
      :  ::  |||||
      389 eLeuYrAsnAsn-----  393
QY      1345 GGATCAATATAAAAATCTGCTGATGAAAGAAATTAACAACACTGACTGATCTTGAGA  1404
      :  ::  |||||  |||||  |||||  |||||
      394 -----His-AsnYrThrAspCySthSerGlu  403
QY      1405 AATATTAGCTCTTAAAAAGGGGAAGAGCCGATATGCTTGTGATGAGAGCTTGA  1464
      403 1y-----  409
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY      1465 AA---CTGTTACCAATCAATAGCTGATGCTGATCAAC---GAATTCG  1509
      :  ::  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      409 ySTrPcySglYrThrGlnAsnYrAspAlaAspGlnlyrPheCySProMeCa  429
QY      1510 TAAAAAGTGAGCACTTTAACAAGCTAGCGAAGCTTAAGCTTCAAGATTATAGC  1569
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
      429 1aAlaHisGlnGluIuLeCySthThrAsnGlu-----  439
QY      1570 ATCCCGTATAGGCTAAAGTAACTACTACAACAAATCGATGCTTTGGTATTAAGACT  1629
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
      440 -----GlyValMeC--T  443
QY      1630 ATACCTTAAGTGAAGAGATATACAGATACACCAACCGTATCATTAACGTTT  1689
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      443 yTArGlnlGluYrAspGlnTrAspYrLysGlnHisAspMetGlyHisMetMeAspCySthrc  463
QY      1690 ATATGCGCAAGCGAAGCCGGAAGAGAGATAGTACCATTTAGCTGCTGGTGGCCAG  1749
      :  ::  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
      463 ySValGlyAsn-----GlyArGlnGlyIu  471
Db      471 rPThrCySthLeuIaYrSerGlnLeuArAspGlnCySthle-----ValAspAspIleT  489
QY      1804 CCTATGAGTGGGAGAAACGTGGAGAGAGCCCTAACAGGCTGATGATGATGATTTGA  1863
      :  ::  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      489 hrTyAsnValAsnAspThrPheHisYrAspGlnHisGlnGluGlnGlnHisMetLeuAsnCyS  509
QY      1864 CTTCCTGGAGAGAGCGAGCGGACATGCTGCACTTTAGAAATAGATGCAAGATC  1923
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
      509 hrCyPheGlnGlnGlnYrAspGlnYrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln  529
QY      1924 AGGAACAAAGACATCTATAGAAATTGGAGACACCTGGAGCAAGAAAGATTAATCGAGAA  1983

```

```

Db      529 eGluThrGlnThrPheTyGlnlGluYrAspSerTrpGluYrValHis---GlyV  548
      ::|||  |||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY      1984 ACCTGCTCCAGTCACTGTCAGACGAGCGGAGAGAGTGGATGAGAGGACGA  2043
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      548 aLArGlyrGlnCySthYrCySthYrGlnYrGlnYrGlnYrGlnYrGlnYrGlnYr  565
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY      2044 CCTGTCGAGACACATGAGCGGAGATCTGGCCCC  2078
      :  ::  |||||  |||||  |||||  |||||  |||||  |||||
Db      566 --ProLeuGlnThrYrProSerSerGlyPro  576
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

RESULT 15
US-10-171-311-64
; Sequence 64, Application US/10171311
; Publication No. US20030087270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Glatc, Karen
; APPLICANT: Gannavarepu, Manjula
; APPLICANT: Hoersh, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; FILE REFERENCE: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; PRIOR FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 64
; LENGTH: 2328
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-171-311-64

Alignment Scores:
Pred. No.: 1,26e-48 Length: 2328
Score: 636.00 Matches: 205
Percent Similarity: 36.56% Conservative: 71
Best Local Similarity: 27.15% Mismatches: 154
Query Match: 17.07% Indels: 326
DB: 4 Gaps: 23

US-09-940-235-12 (1-2096) x US-10-171-311-64 (1-2328)
QY      5 GACCATTCATGTTGTTGCTCTAGGTCGACAGCGTTTGCACAGAGTCGCTTCACT  61
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      95 AspCySthrcYsllGlnYrGlnYrGlnYrGlnYrGlnYrGlnYrGlnYrGlnYr  114
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY      62 -----TCGCTCGGTAATCGGTGATTCCTTCTGCTAACCAATAGAGCAACC  106
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      115 HisGlnGlnYrGlnYrGlnYrGlnYrGlnYrGlnYrGlnYrGlnYrGlnYrGlnYr  130
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY      107 CCGCAGCCTTACGCGGGGTCCTCAACGACAGACAGCATGCGCACCCGTGGCCAGA  166
      :  ::  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      130 oHisGlnYr-----  133
QY      167 CCCAAGCGTCCCGAGATCTGATCCGCGAAATTAATAGACTACTATAGGAGACGA  226
      133 -----
Db      133 -----  133
QY      227 CAACGGTTCCCTTGAAG---ATAATTTGTTTAACTTTAAGAAAGATATACAGAG  283
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
      134 -GlyGlnYrMetLeuGlnCySthValCySthGlnYrGlnYrGlnYrGlnYrGlnYr  151

```


GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 28, 2006, 02:09:11 ; Search time 6.06893 Seconds
(without alignments)
7479.985 Million cell updates/sec

Title: US-09-940-235-12
Perfect score: 3726
Sequence: 1 cgaagaccatcatgtgtgtt.....cctcacgcgtgtcgttag 2096

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 75621 seqs, 10829074 residues
Total number of hits satisfying chosen parameters: 151242

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -MODEL=frame+ n2p.model -DEV=xlp
-O=/cgn2_1/USPTO_epool_p/US09940235/runat_27012006_144220_27676/app_query.fasta_1.7708
-DB=Published Applications_AA_New -QFMT=faetan -SUFFIX=rapbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pro -NORM=ext -HEARSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09940235_@CGN_1_1_@runat_27012006_144220_27676
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications_AA_New:*
1: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubppaa/PCT_NEW_PUB.pep:*
5: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pep:*
7: /cgn2_6/ptodata/2/pubppaa/US11_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	1984	53.2	414	US-10-631-558-2
2	1928	51.7	413	US-11-032-951-12
3	634.5	17.0	657	US-10-995-561-622
4	634.5	17.0	984	US-10-995-561-629
5	634.5	17.0	2296	US-10-995-561-633
6	634.5	17.0	2355	US-10-995-561-623
7	634.5	17.0	2355	US-10-995-561-627
8	634.5	17.0	2384	US-10-821-234-1545
9	634.5	17.0	2386	US-10-995-561-626
10	632	17.0	259	US-10-631-558-4

11	632	17.0	300	US-11-006-119-31	Sequence 31, App1
12	632	17.0	642	US-10-995-561-631	Sequence 621, App
13	278.5	7.5	1341	US-10-995-561-621	Sequence 631, App
14	275.5	7.4	693	US-10-995-561-632	Sequence 632, App
15	275.5	7.4	1315	US-10-995-561-630	Sequence 630, App
16	275.5	7.4	1348	US-10-995-561-624	Sequence 624, App
17	260	7.0	1259	US-10-995-561-625	Sequence 625, App
18	221	5.9	1286	US-10-995-561-628	Sequence 628, App
19	202.5	5.4	847	US-10-995-561-634	Sequence 634, App
20	132.5	3.6	1992	US-11-013-759-3	Sequence 3, App1
21	132.5	3.6	1992	US-11-013-759-13	Sequence 13, App1
22	132.5	3.6	2047	US-11-013-759-4	Sequence 4, App1
23	132.5	3.6	2047	US-11-013-759-7	Sequence 7, App1
24	125	3.4	1417	US-11-052-554A-8	Sequence 8, App1
25	119	3.2	2340	US-11-052-554A-171	Sequence 171, App
26	116.5	3.1	470	US-10-485-517-319	Sequence 319, App
27	116.5	3.1	1647	US-11-052-554A-260	Sequence 260, App
28	115.5	3.1	2890	US-11-115-639-31	Sequence 31, App1
29	115.5	3.1	2890	US-11-115-639-32	Sequence 32, App1
30	115.5	3.1	2890	US-11-115-639-33	Sequence 33, App1
31	115	3.1	21	US-11-032-951-7	Sequence 7, App1
32	115	3.1	463	US-10-793-626-960	Sequence 960, App
33	113	3.0	801	US-11-174-150-29	Sequence 29, App1
34	113	3.0	801	US-11-124-368A-292	Sequence 292, App
35	113	3.0	1765	US-10-055-877-140	Sequence 140, App
36	112.5	3.0	1940	US-10-055-877-141	Sequence 141, App
37	112	3.0	408	US-11-016-564-5	Sequence 5, App1
38	110.5	3.0	1448	US-10-485-517-212	Sequence 212, App
39	108.5	2.9	2314	US-11-013-759-11	Sequence 11, App1
40	108	2.9	4913	US-10-453-372-1142	Sequence 1142, App
41	108	2.9	4961	US-10-453-372-1132	Sequence 1132, App
42	107.5	2.9	948	US-10-485-517-131	Sequence 131, App
43	107	2.9	1416	US-11-128-059-60	Sequence 60, App1
44	107	2.9	1494	US-11-128-059-78	Sequence 78, App1
45	107	2.9	2086	US-11-128-059-82	Sequence 82, App1

ALIGNMENTS

RESULT 1
US-10-631-558-2
; Sequence 2, Application US/10631558
; Publication No. US20050260598A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kamara
; APPLICANT: Mahalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/10/631,558
; PRIOR FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US/09/940,235
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Streptococcus equisimilis
US-10-631-558-2
Alignment Scores:

Pred. No.: 1.52e-141 Length: 414
 Score: 1984.00 Matches: 383
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 53.25% Indels: 0
 DB: 6 Gaps: 0

US-09-940-235-12 (1-2096) x US-10-631-558-2 (1-414)

```

QY 588 ATTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACGCAATTGTTGTT 647
Db 1 IleaIaGIyProGIuTrpLeuLeuAspArgProSerValAsnAsnSerGIuLeuVal 20
QY 648 AGCGTTGCTGCTGTTGAGGGGAGCAATTCAGACATTAAGTCTTAATTTTGAATC 707
Db 21 SerValaIaGIyThrValaGIuGIyThrAsnGIuAspIleSerLeuysPhepNegIuIle 40
QY 708 GATCTAACATCAAGCACTGCTCATGAGGAGAAAGACAGCAAGGCTTAAGTCCAAATCA 767
Db 41 AspLeuThrSerArgProIaIaHISGIyGIyLysThrGIuGIuLysLeuSerProLysSer 60
QY 768 AAACCATTTGCTACTGATAGTGGCGGATGTGCATAACTTGAGAAAGCTGACTTACTA 827
Db 61 LysProPheAlaThrAspSerGIyAlaMetSerHisLysLeuGIuLysAlaAspLeuLeu 80
QY 828 AAGGCTATTCAAGAACATTTGATGCTAGCGTCCACAGTAAGACGACATCTTGAGGTC 887
Db 81 LysAlaIleGIuGIuGIuGIuLeuIleAlaAsnValHisSerAsnAspAspIyrPheGIuVal 100
QY 888 ATTGATTTTGGACGAGATGACCACTTAATCTGCAAGGCAAGGCTACTTGTGTCAC 947
Db 101 IleAspPheAlaThrAspAlaThrIleThrAspAspAsnGIyLysValIyrPheAlaAsp 120
QY 948 AAAGATGTTGCTGTAACCTTGCGGACCCACCTGTCCAGAAATTTTGTAAAGCGGACAT 1007
Db 121 LysAspGIySerValThrLeuProThrGIuProValGIuGIuPheLeuLeuSerGIyHis 140
QY 1008 GTGGCGGTTTAAACCATTAATAAAGAACCAATATCAAAACAGCGCAATCTGTATGTCG 1067
Db 141 ValArgValArgProLysGIyLysProIleGIuAsnGIuAlaLysSerValaAspVal 160
QY 1068 GAATTAATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTGACAGGTCGAA 1127
Db 161 GIuYrThrValGIuPheThrProLeuAsnProAspAspPheAspProGIyLeuLys 180
QY 1128 GATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTACTA 1187
Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGIyAspThrIleThrSerGIuGIuLeuLeu 200
QY 1188 GCTCAAGACAAAGATTTTAAACAAACCAACCCAGGCTATACGATTTATGAACGTGAC 1247
Db 201 AlaGIuAlaGIuGIuSerIleLeuAsnLysAsnHisProGIyLyrThrIleLyrGIuAsp 220
QY 1248 TCCTCAATCGTCACTGACATGACAAATGATTTTCCGATAGATTTTACATGATGATCAAG 1307
Db 221 SerSerIleValIThrHisAspAsnAspIlePheArgThrIleLeuProMetCspGIuGIu 240
QY 1308 TTTTACTACCGTGTAAAAATCGGAGACAGCTTATAGGATCAATAAAAATCTGCTGTC 1367
Db 241 PheThrLyrArgValLysAsnArgGIuGIuAlaLyrArgIleAsnLysLysSerGIyLeu 260
QY 1368 AATGAGAAATTAACAACATGACCTGATCTTGAGAAATTAATTCGCTTTAAAAAGG 1427
Db 261 AsnGIuGIuLysLysAsnThrAspLeuLysSerGIyLysLyrValLeuLysLysGIy 280
QY 1428 GAAAGACGATAGATCCCTTGTGATGACGATCACTTGAAACGTGTCAACGCAATCGTT 1487
Db 281 GIuLysProLyrAspProPheAspAspSerHisLysLysLeuPheThrIleLysLyrVal 300
QY 1488 GATGTCATACCAACGAATTTGCTAAAAAAGTGACAGCTTTAAACAGCTACGCAACGTAAC 1547
Db 301 AspValaAspThrAsnGIuLeuLeuLysSerGIuGIuLeuLeuThrLysSerGIuArgAsn 320

```

```

QY 1548 TTAGACTTCAGAGATTTATACGATCTCGTATAGGCTTAAAGCTTACTTACAAATCTC 1607
Db 321 LeuAspPheArgAspLeuLyrAspProAspLysAlaLysLysLeuLyrAsnAsnLeu 340
QY 1608 GATGCTTTTGGTATTATGACATTAACCTTAACCTGAAAGATAGAGATTAATCAAGTAC 1667
Db 341 AspAlaPheGIyLysLeuMetAspLyrThrLeuThrGIyLysValGIuAspAsnHisAspAsp 360
QY 1668 ACCAAGCGTATCAATACCGTTTATATGGCGAAGCGACCCGAGAGAGAAATGCTAGTAC 1727
Db 361 ThrAsnArgIleIleThrValIyrMetGIyLysArgProGIuGIuAsnAlaSerLyr 380
QY 1728 CATTTAGCT 1736
Db 381 HisLeuAla 383

```

RESULT 2

```

US-11-032-951-12
; Sequence 12, Application US/11032951
; Publication No. US20050282741A1
; GENERAL INFORMATION:
; APPLICANT: Kyotal, Gerald
; APPLICANT: Rabkin, Simon W.
; TITLE OF INVENTION: Peptides and Their Use to Ameliorate
; TITLE OF INVENTION: Cell Death
; FILE REFERENCE: 50216/003005
; CURRENT APPLICATION NUMBER: US/11/032,951
; PRIOR FILING DATE: 2005-01-11
; PRIOR APPLICATION NUMBER: US 09/919,703
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 09/294,457
; PRIOR FILING DATE: 1999-04-19
; PRIOR APPLICATION NUMBER: US 08/759,599
; PRIOR FILING DATE: 1996-12-05
; PRIOR APPLICATION NUMBER: US 60/008,233
; PRIOR FILING DATE: 1995-12-06
; NUMBER OF SEQ. ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 413
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic polypeptide
US-11-032-951-12

```

Alignment Scores:

```

Pred. No.: 2.38e-137 Length: 413
Score: 1928.00 Matches: 377
Percent Similarity: 98.95% Conservative: 1
Best Local Similarity: 98.69% Mismatches: 2
Query Match: 51.74% Indels: 2
DB: 7 Gaps: 2

```

US-09-940-235-12 (1-2096) x US-11-032-951-12 (1-413)

```

QY 588 ATTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACGCAATTGTTGTT 647
Db 1 IleaIaGIyProGIuTrpLeuLeuAspArgProSerValAsnAsnSerGIuLeuVal 20
QY 648 AGCGTTGCTGCTGTTGAGGGGAGCAATTCAGACATTAAGTCTTAATTTTGAATC 707
Db 21 SerValaIaGIyThrValaGIyThrAsnGIuAspIleSerLeuysPhepNegIuIle 39
QY 708 GATCTAACATCAAGCACTGCTCATGAGGAGAAAGACAGCAAGGCTTAAGTCCAAATCA 767
Db 40 AspLeuThrSerArgProIaIaHISGIyGIyLysThrGIuGIuLysLeuSerProLysSer 59
QY 768 AAACCATTTGCTACTGATAGTGGCGGATGTGCATAACTTGAGAAAGCTGACTTACTA 827
Db 60 LysProPheAlaThrAspSerGIyAlaMetSerHisLysLeuGIuLysAlaAspLeuLeu 79
QY 828 AAGGCTATTCAAGAACATTTGATGCTTAACGTTCAAGTAAGCAAGCACTTTGAGGTC 887

```

```

Db      |||
      80  LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspArgTyrPheGlnVal 99
Qy      888  ATGTATTTTGGCAAGCGATGCAACCATTTATGATCGAAACGGCAAGGTCTATTGTCTGAC 947
Db      100  IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyValTyrPheAlaAsp 119
Qy      948  AAAAGATGTTCCGTTAACCTTCCGCAACCTGTCTCCAGAAATTTTGTCTAACGGCAAT 1007
Db      120  LysAspGlySerValThrLeuProThrGlnProValGlnGlnPheLeuLeuSerGlyHis 139
Qy      1008  GTGGCGCTTAGACCATTTAAAGAAAAACAATAACAACCAACGGAATCTGTGATGTG 1067
Db      140  ValArgValArgProTyrLysGlnLysProIleGlnAsnGlnAlaLysSerValAspVal 159
Qy      1068  GAATATACTGTACAGTTTATCCCTTAAACCTGTATGACAGATTTTCAGACAGGCTCAAA 1127
Db      160  GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 179
Qy      1128  GATTAAGCTATTTGAAAAACAATGATGTGGTGACACCATCAATCTCAAGAAATTAATA 1187
Db      180  LeuThrLysLeuLeuLysThrIleValIleGlyAspThrIleThrSerGlnGlnLeu 199
Qy      1188  GCTCAAGCAACAAGCATTTTAAACAACCAACCGGCTATACGATTTTGAACGTGAC 1247
Db      200  AlAGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGlnArgAsp 219
Qy      1248  TCCTCAATGCTCATCATGACATGACATTTTCCGTAACGATTTTAAACAAGTCAAGAG 1307
Db      220  SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGln 239
Qy      1308  TTTATCTTACCGTGTAAAAATCGGGAACAAGCTTATAGATCAATAAAAAATCGTGTG 1367
Db      240  PheThrTyrArgValLysAsnArgGlnGlnAlaIleTyrArgIleAsnLysLysSerGlyLeu 259
Qy      1368  AATGAAGAAATTAACAACATGACCTGATCTCTGAGAAATAT--TACGTCTTAAAAAA 1424
Db      260  AsnGlnGlnIleAsnAsnThrAspLeuIleSerLeuGlnLysTyrValLysLys 279
Qy      1425  GGGGAAAGCCGATATGATCCCTTGTATGCGAGTCACTTGAATGTTTCAACCAATAC 1484
Db      280  GlyGlnLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyr 299
Qy      1485  GTTGAATGTCATTAACAAGAAATGTCTAAAGAGAGAGCTTAAACAGCTAGCGAAGCT 1544
Db      300  ValAspValAspThrAsnGlnLeuLeuLysSerGlnGlnLeuThrAlaSerGlnArg 319
Qy      1545  AACTTAGACTTCAAGATTTATACGATCTCGTATAGGCTTAACTACTCTTCAACAAT 1604
Db      320  AsnLeuAspPheArgSerLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsn 339
Qy      1605  CTCGATGCTTTGGTATTATGCACTTACCTTACTGAAAGATGAGATTAATCAAGAT 1664
Db      340  LeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyValGlnAspAsnHisAsp 359
Qy      1665  GACACAACCGTATCAACCGTTTATATGAGCAACGACCGCAAGAGAGAAATGCTAGC 1724
Db      360  AspThrAsnArgIleIleThrValTyrMetGlyLysArgProGlnGlnAlaAlaSer 379
Qy      1725  TACCAT 1730
Db      380  TyrHis 381

```

```

; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 622
; LENGTH: 657
; TYPE: PRN
; ORGANISM: Homo sapiens
US-10-995-561-622

Alignment Scores:
Pred. No.: 1,94e-40 Length: 657
Score: 634.50 Matches: 207
Percent Similarity: 36.69% Conservative: 70
Best Local Similarity: 27.42% Mismatches: 151
Query Match: 17.03% Indels: 328
DB: Gaps: 28

US-09-940-235-12 (1-2096) x US-10-995-561-622 (1-657)
Qy      5  GACCATCATGTTGTGTCAGGTGGAGAGCTTTTGACAGAGAGTGGCTTCAAGT--- 61
Db      122  AspCysThrCysIleGlyAlaGlyArgGlyArgIleSerCysThrIleAlaAsnArgCys 141
Qy      62  -----TGCCTGCGTATCGATGATTCATTTCTGTAAACAGTAAGCAAC 106
Db      142  HisGlnGlyGlnGlnSerTyr-LysIleGlyAspThrTrp-----ArgArgPr 157
Qy      107  CGGCCAGCCTAGCCGGGCTCTCAACAGACAGACATCATGCGCACCCGTGGCCAGGA 166
Db      157  ohIeGlnThr----- 160
Qy      167  CCCAAGCCTGCCAGATCTCGATCCCGGCAATTAATAGATCTATAGGAGACCA 226
Db      160  ----- 160
Qy      227  CAACGGTTCCCTAGAA---ATAATTTGTTTAACTTTAAGAGAGATATACATCG 283
Db      161  -GlyGlyTyrMetLeuGlnCysValCysLeuGlyAsnGlyLysGlnGln-TrpThr--- 178
Qy      284  TGCAAGCAACAAGATTTGTAACCATAGCTGAGAGATGTTTGTATCATGCTGCGGACTT 343
Db      179  -----CysLysProIleAlaGlnLysCysPheAspHisAlaIleGlyThrS 194
Qy      344  CTTATGTCGTCGAGAAACGTGGA----- 368
Db      194  eTyrValValGlyGlnThrTrpGlnLysProTyrGlnGlyTyrTrpMetValAspCys 214
Qy      369  -----GAAGCAGGAGGAGCGCATCTTGCACTTGTAGAAATAGATGCAAGATC 417
Db      214  hrcysleuGlyGlnGlySerGlyArgIleThrCysThrSerArgAsnArgCysAsnAspG 234
Qy      418  AGGACAACAAGACATCTTATAGATTTGAGACACCTGAGACCAAGAGATATGAGAAA 477
Db      234  InAspThrArgThrSerTyrArgIleGlyAspThrTrpSerLysLysAspAsnArgGly 254
Qy      478  ACTGCTCCAGTCATGTGACAGCAAGCAAGCGCCAGGAGAGATGAGATGAGAGCA 537
Db      254  snleuLeuGlnCysIleCysThrGlyAsnGlyArgGlyGlnTrpLysCysGlnArgHis 274
Qy      538  CTTCTGTGAGACACATCGAGGGGATCTGGCCCTTCAACCGATGTCGATTTGCT--- 593
Db      274  hrcervalGlnThrThrSerSerGlySerGlyProPheThrAspValArgAlaVal 294
Qy      593  ----- 593
Db      294  yrcInProGlnProHisProGlnProProProTyrGlnHisCysValAlaAspSerGly 314
Qy      594  -----GACCTGAGTGGCTGTAGACCGTCCATCTGTGCAACAACGCAAT 639
Db      314  alValTyrSerValGlyMetGlnTrpLeu-----LysThrGlnGlyAsnLysGln 331
Qy      640  TGGTTGTTAGCGTTGCTGTGACTGTTGAGGGGAGCAATCAAGACATTAATGCTTTT 699

```

```

RESULT 3
US-10-995-561-622
; Sequence 622, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01559

```

Db	331	etleuCythrCylleuclY-----	337
Oy	700	TTGAATGCATCTTAACATCAAGCACTGCTCATGAGGAAAGACAGAGCGATTAACTC	759
Db	338	-----	341
Oy	760	CAAAATCAAAACCAATTTCGTACTGATAGTGGCGCAATGTCACATAACTTAGAAAGCTG	819
Db	341	-----	341
Oy	820	ACTTACTAAAGCCTAATTCAGAAACAATTGATCGTCAACGTCACAGTAAACGACACTAC	879
Db	342	-----	357
Oy	880	TTGAG-----GTCAATATTTTTCAGAGGATGCAACCAATTACGTATGAAACGGCAAG	933
Db	357	lygluProCyvalleuProPheThr-----	370
Oy	934	TCCTACTTTCGCT-----	978
Db	370	hrPheTySerCythrThrIgluIyAArgInAspIyHileuTyPcySerThrThrs	390
Oy	979	CTGTCCAGAAATTTTGTGTAAGCGGACACATGTCGGGTTAAGACCAATATAAGAAACCA	1038
Db	390	erAsnTyGlu-----	393
Oy	1039	TACAAACCAAGCGAAATCTGTGTGATGTGCAATATATCTGTACAGTTTACTCCCTTAAAC	1098
Db	394	--GlnAspGlnIlySerIySerPheCythrIAspHileThrVal-----	406
Oy	1099	CTGATGACGATTTACAGACACAGGCTCAAAAGATCTAAGCTATTGAAACACTAGCTATCG	1158
Db	406	-----	406
Oy	1159	GTGACACCATGACATCTGAAGATTTACAGTCTCAAGCAACAAAGATTTTAAACAAAAC	1218
Db	407	-----	412
Oy	1219	ACCAGGCTATACGATTTATGCAACGTGACTCTCTCAATCGTCACTGATGACAATGACATTT	1278
Db	413	-----	421
Oy	1279	TCCGTACGATTTTACCAATGATCAAGAGTTTACTTACCGTGTAAAAATCGGAAACAG	1338
Db	422	-----	429
Oy	1339	CTTATAGCATCAATTAATAATCTGGTCTGATGAAAGAAATAAACCAACCTGATCT	1398
Db	430	-----	436
Oy	1399	CTGAGAAATATTACGCTCTTAAAAAAGGGGAAAAAGCGATGATGCTCTTGATGCGCAGTC	1458
Db	436	erGluIyGly-----	442
Oy	1459	ACTTGAAA--CTGTTCACCATCAAAATACGTTGATGCGATACCAAC-----	1503
Db	442	smleTyIytrCyseIyThrThrGlnIAsnTyIyAspAlaAspGlnIlyAspGlyPheCySP	462
Oy	1504	AATTGCTAAAAAGTACAGACGCTTTAAACAGCTGACGACAGTAACTTAAGACTTCAGAGATT	1563
Db	462	romelAlaAlaIleIleGluIuIleCythrThrAsnGlu-----	474
Oy	1564	TATACGATCTCGTGTATAAAGGTAAAGCTAACTGTATACAAATCTGATGCTTTGGTATTA	1622
Db	475	-----	477
Oy	1624	TGCACTATACCTTAACTGAAAAAGTAGAGATTAATCAAGATGACCAACCGTATCATATA	1683
Db	477	et---TyAArgIleIleIyAspGlnItrPArgIlyGlnIleAspMetCylIyIleMetMetAArgC	496
Oy	1684	CCGTTTATATGGGACAGCGAACCCGGAAGAGAGATGTACTGATCAACTTTAGCTGTGGTG	1743

```

Db      496  ySthCyValGlYshn-----GlyArgS 504
QY      1744  GCGAGGCGGACAGATTGTA-----CCCATAGCTGAGAAAGTGTTTATCATGCTGCTG 179
Db      504  lYglUtrPthrCySlleatYrSergInleuArgspInCysIle-----ValAspA 522
QY      1798  GGACTTCTCATATGCTGGAGAAAGCTGGGAGAAAGCCCTAACAAAGCTGGATGATGTTAG 185
Db      522  spIethrTyxsnValAsnAspRtrPheNlVlyAsrIghIsgIugIyNlMetleuA 542
QY      1858  ATTGACTTCTGCTGGAGAGAGCGGAGGAGCATCTGACTTCTTAGAAATGATGCA 191
Db      542  enCyThrCyehedIyngInglYArgIyArTrIySsAspRProValAspInCysG 562
QY      1918  ACGATCAGACACACAGACATCTTATAGATTGGAGACACCTGGACACAGAGAGATATC 197
Db      562  lAspserdIuThrIghIyThrPheTyrgInIleIyAspserTrpIuIySyTyValhIs 581
QY      1978  GAGGAACCTGCTCCAGTGCATCTGCACAGGACGCGCGAGAGAGATGTAATGTGACA 203
Db      582  --GlyValArGrTyrgInCysTyTyGlyArGdIyIleIyGluTrpNhIsCyseIn 600
QY      2038  GGCACACCTCTGTGAGACACCATGCAGGCGGATCTGACCC 2078
Db      601  -----ProleuGInThrTyTrProserSerSergIyPro 611

RESULT 4
US-10-995-561-629
; Sequence 629, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 629
; LENGTH: 984
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-629

Alignment Scores:
Pred. No.: 2,1e-40 length: 984
Score: 634.50 Matches: 207
Percent Similarity: 36.69% Conservative: 70
Best Local Similarity: 27.42% Mismatches: 151
Query Match: 17.03% Indels: 328
Db: 6 Gaps: 28

US-09-940-235-12 (1-2096) x US-10-995-561-629 (1-984)
QY      5  GACCATTCATGTTGTTGCTCAGAGTCGACAGACGTTTTCAGACAGATCGCTTCACT--- 61
Db      122  AspCyThrNrySlleGlyAlaheIyArGdIyArgIleSerCyehrIleAlaAnaYgys 141
QY      62  -----TCGCTCGCGTATCGGTATTCATTCCTGCTAACAGTAAGCAACC 106
Db      142  HisIugIyGlyGInserTyr-LysIleGlyAspTrIrp-----ArgArGr 157
QY      107  CGCGACGCTAGCCGGGCTCTCAAGACACAGAGACAGATCATGCGACCCGTGGCAGGA 166
Db      157  OhIsGluThr----- 160
QY      167  CCNAAGCTGCCGAGATCTGCATCCCGGAATTAATACACTCACTAATAGGAGACGA 226
Db      160  ----- 160
QY      227  CAACGGTTCCCTCTAGAA---ATAATTGTTGTTAACTTTAAGAGAGATATACATGG 283

```



```

; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 633
; LENGTH: 2296
;
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-995-561-633

```

Alignment Scores:	
Pred. No.:	2 466-40
Score:	634.50
Percent Similarity:	38.69%
Best Local Similarity:	27.42%
Query Match:	17.03%
DB:	6
	Gaps: 28
	Length: 2286
	Matches: 207
	Conservative: 70
	Mismatches: 151
	Indels: 328
	Gaps: 28

US-09-940-235-12 (1-2096) x US-10-995-561-633 (1-2296)

QY	5	GACCACTTCAATGTTGTTGCTCAGGTGGCAGACGTTTGGACAGACATGCGTTTACGTT---	61
Db	122	AspCysThrCysAlIeGlyAlaGlyArgGlyIleSerCysThrIleAlaIaArgCys	141
QY	62	-----TCGCTCGGATTCGGTGAATTCATCTGTCTAACCAAGTAAGGCAACC	106
Db	142	HisGluGlyGlyIleGlnSerTyr-LysIleGlyAspMtrTP-----ArgArgPr	157
QY	107	CCGCAGCGCTAGCCGGGTCCTCAACAGCAGAGACAGATCATGCGCACCCGGTGGCAGGA	166
Db	157	OHlIeGluThr-----	166
QY	167	CCCAACGCTGCCGAGATCTCGATCCCGGAAATTATACACTCTATTAGGAGACCA	226
Db	160	-----	160
QY	227	CAAGCGTTTCCCTTAGAA---ATAATTGTTTACTTTAGAAGAGATATACCATGG	283
Db	161	-GlyGlyTyrMetLeuGluCysValCysLeuGlyAsnGlyLysGlyGlu-TripThr---	178
QY	284	TGCAAGCACAACAGATTGACCATGATGACTGAGAAAGTTTGTATCATGCTGCTGGACCTT	343
Db	179	-----CysLysProIleAlaGluLysCysPheAspHisAlaAlaGlyThrS	194
QY	344	CCTATGTGCTCGAGAAACGTGGAA-----	368
Db	194	erTyrValValGlyGluThrTrpGluLysProTyrGlnGlyTrpMetValAspCysT	214
QY	369	-----GAAGCAGCGGACCGATTCATCTTGCACTTCTTAGAAATAGATGCAAGATC	417
Db	214	hrCysLeuGlyAlaGluGlySerGlyArgGlyIleThrCysThrSerArgAsnArgCysAsnAspG	234
QY	418	AGGACACAGAGACATCTCTTAGAATTGAGAGACACCTGGAGCAGAAAGAGATTAATCGAGAA	477
Db	234	lnAspThrArgThrSerTyrArgGlyIleGlyAspThrTrpSerLysAspAsnArgGlyA	254
QY	478	ACCTGCTTCAGTGCATCTTGCACAGGCAAGCGCCGAGAGAGTGGAGTGGAGGACACA	537
Db	254	snLeuLeuGlnCysIleCysThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisT	274
QY	538	CCTCTGTGCAAGCACATGTGAGCGGATCTGGCCCTTACCCGATGTTCCGATTGCT---	593
Db	274	hrSerValGlnThrThrSerSerGlySerGlyProPheThrAspValArgAlaAlaValT	294
QY	593	-----	593
Db	294	yrGlnProGlnProHisProGlnProProIleTyrGlyHisCysValThrAspSerGlyV	314
QY	594	-----GACCTGAGTGGCTGCTAGACCGTCACTTCATGTGCAACAAGCAACAT	639
Db	314	alValTyrSerValGlyMetGlnTrpLeu-----LysThrGlnGlyAsnLysGluM	331
QY	640	TGCTTTAGCGCTTCTGGTACTGTGTTGAGGGAGCAATCAAGACATTAAGCTTTAAATTTT	699
		::::: :::	

Db	331	eLeuCy ^{thr} Cy ^{leu} Leu ^{ly} -----	337
Qy	700	TTGAATGCATCTTAACATCAGCACTGCTCATGAGGAAAGACAGAGGCTTAACTC	759
Db	338	-----	341
Qy	760	CAAAATCAAAACATTGCTACTGATTAATGCGCGGATGTCATATACTTGAGAAAGCTG	819
Db	341	-----	341
Qy	820	ACTTACTAAAGGCTATTGACGAAACAATTGATCGCTAAAGTCAACGTAACGACGACTACT	879
Db	342	-----	357
Qy	880	TTGAG-----GTCAATTGATTTTGGCAAGCGATGCAACCATTAAGTATGATCGAAACGGCAAG	933
Db	357	LyGluProCy ^{val} LeuPro ^{thr} Thr-----	370
Qy	934	TCCTACTTTGCT-----	978
Db	370	hrPheTy ^{ser} Cy ^{thr} Thr ^{thr} Glu ^{glu} Arg ^{glu} Asp ^{glu} Leu ^{thr} Cy ^{ser} Thr ^{thr} Thr ^{thr} S	390
Qy	979	CTGTCCAGAAATTTTGGTAAAGCGGACATGCGCGTTAAAGACATATAAGAAAAACCA	1038
Db	390	erA ^{thr} Tr ^{glu} -----	393
Qy	1039	TACAAACCAAGCGAAATCTGTGTGATGTGAATATACTGTACAGTTTACTCCCTTAAAC	1098
Db	394	--GlnAspGln ^{ly} Ser ^{tyr} Phe ^{tyr} Asp ^{thr} Asp ^{thr} Val-----	406
Qy	1099	CTGATGACGATTTTGACACACAGGCTCAAAAGATCAATAGCTATGTAACACTAGCTATCG	1158
Db	406	-----	406
Qy	1159	GTGACACCATGACATCTCAAGAATTAATAGTCAAGCAACAAGCATTTTAAACAAAAC	1218
Db	407	-----	412
Qy	1219	ACCAGGCTATAGCATTTATGACAGCGTCACTCTCAATAGTCAATCATGACATGACATGATTT	1278
Db	413	-----	421
Qy	1279	TCCGTACGATTTTACCAATGATGACAGAGTTTACTTACCGTGTAAAAATCGGAAACAG	1338
Db	422	-----	429
Qy	1339	CTTATAGCATCAATATAAAAATCTGCTGATGTAAGAAATTAACAACACTGACCTGATCT	1398
Db	430	-----	436
Qy	1399	CTGAGAAATATTACGTCCTTAAAAAGGGGAAAAAGCGTATGATCCCTTGTATGCGAGTC	1458
Db	436	erGlu ^{glu} Ly-----	442
Qy	1459	ACTTGAAG--CTGTCAACATCAAAATACGTGTATGTGATACCAAC-----G	1503
Db	442	smMet ^{ly} TrpCy ^{glu} Ly ^{thr} Thr ^{glu} Asn ^{tyr} Leu ^{tyr} Ala ^{asp} Gln ^{ly} Ser ^{glu} Leu ^{tyr} Phe ^{Cys} Pro	462
Qy	1504	AATTGCTAAAAAGTACGACGCTCTTAAACAGCTAGACGAACGTAACCTTAGACTTCAGACATT	1563
Db	462	rometAla ^{ala} Leu ^{ala} Gln ^{glu} Leu ^{thr} Leu ^{tyr} Thr ^{thr} Asn ^{glu} -----	474
Qy	1564	TATAGCATCTCGGATAGTAAGCTTAAAGCTAAACATCTCAACAACATCTGATGCTTTTGGTATTA	1623
Db	475	-----	477
Qy	1624	TGCACTATACCTTAACTGAAAAGTACAGATTAATCAAGATGACACCAACCGTATCATTA	1683
Db	477	et-----Ty ^{arg} Gln ^{leu} Leu ^{asp} Gln ^{thr} Trp ^{ser} Leu ^{glu} Asp ^{met} Glu ^{tyr} Asn ^{thr} Met ^{thr} Ala ^{arg} C	496
Qy	1684	CCGTTTATATGGGCAAGGACCCGAGAGAGAAATGCTAGCTACCATTTAGCTGGTGTG	1743
Db	496	y ^{thr} Trp ^{Val} Gln ^{glu} Asn-----	504

```

Qy 1744 GCCAGCGCAGACAGATTGTA-----CCCATACCTGAGAGTGTTCATGATCAGTCTGTC 1797
Db 504 lYgluTrpThrCyslleAlaTyrsErGlLeuArGsrInGlySile-----ValAspA 522
Qy 1798 GGACTTCCATATGTGTGCGAGAAACGTGGAGAACCCCTACCAAGCGTGGATGATGTTAG 1857
Db 522 spIeThrTyAsnValAsnAspThrPheHieTyAsrGhIsGluGluGlyHieMetLeuA 542
Qy 1858 ATTGACTTGTGCTGGAGAGGAGCGAGCGCATCTGACCTTCTAGAAATAGATGCA 1917
Db 542 sncYthrCysPheGlyGlnGlyArgGlyArGTrpLysCysAspProValAspGlnCysG 562
Qy 1918 ACGATCAGACACAAGACATCTTATGATTAATGAGACACCTGACGACAGAAAGATATGAC 1977
Db 562 lAspSerGluThrGlyThrPheTyrcInlIeGlyAspSerTrpGluTyTyValHis- 581
Qy 1978 GAGGAAACCTGCTCCAGTGCATCTGCACAGGCAACGCGCGAGAGAGTGGAACTGTAGA 2037
Db 582 --GlyValArGTrpGlnCysTyrcTyrcTyrcTyrcTyrcTyrcTyrcTyrcTyrcTyrc 600
Qy 2038 GGCACACCTGTGCGACACATGAGCGGAGATCTGGCCCC 2078
Db 601 -----ProLeuGlnTrpTyrcProSerSerGlyPro 611

RESULT 6
US-10-995-561-623
; Sequence 623, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 623
; LENGTH: 2355
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-623

Alignment Scores:
Pred. No.: 2,486-40 Length: 2355
Score: 634.50 Matches: 207
Percent Similarity: 36.69% Conservative: 70
Best Local Similarity: 27.42% Mismatches: 151
Query Match: 17.03% Indels: 328
Gaps: 28

US-09-940-235-12 (1-2096) x US-10-995-561-623 (1-2355)
Qy 5 GACCATTAATGTTGTCAGAGTGGACAGCGTTTGACAGACAGTCCGTTACGT--- 61
Db 122 AspCysThrCysIleGlyAlaGlyArgGlyArgGlyIleSerCysThrIleAlaAsnArgCys 141
Qy 62 -----TCGCTCGATATCGGTATGATTCATCTGCTAACAAGTAAAGCAACC 106
Db 142 HisGluGlyGlnIleSerTyrc-LysIleGlyAspTrp-----ArgArGTr 157
Qy 107 CCGCAGACCTTACCGCGGTCTCTCAACAGACAGACAGATACGACACCGTGGCCAGGA 166
Db 157 ohIeGluThr----- 160
Qy 167 CCGAAGCTGCCAGATCTCGATCCCGGAAATTAATACACTACTATAGGAGAGCA 226
Db 160 ----- 160
Qy 227 CAAGGTTTCCTCTAGAA--ATAATTTTGTTTAACTTTAAGAGAGATATACATG 283

```

```

Db 161 -GlyGlyTyrcMetLeuGlnCysValCysLeuGlyAsnGlyLysGlyGlu-TrpThr----- 178
Qy 284 TGCAAGCAACAAGATTGTATCCCATAGCTGAGAGTGTGTTGATCATACCTGCGACTT 343
Db 179 -----CysLysProIleAlaGlnLysCysPheAspHieAlaGlyHis 194
Qy 344 CTAATGTGTCGAGAAACGTGGA----- 368
Db 194 erTyrcValGlyGlnThrTrpLysProTyrcGlnGlyTyrcMetMetValAspCysT 214
Qy 369 -----GAGCGAGCGGACGATCACTTGCATCTTCTAGAAATAGATGACATG 417
Db 214 hrCysLeuGlyGlnGlySerGlyArGTrpGlyLeThrCysThrSerArGAsnArGTrpCysAsnArG 234
Qy 418 AGGACACAGAGACATCTTATGATTAATGAGACACCTGAGAGAAAGATTAATGAGAA 477
Db 234 lAspTrpArGThrSerTyrcTyrcTyrcTyrcTyrcTyrcTyrcTyrcTyrcTyrcTyrcTyrc 254
Qy 478 ACCTGCTCCAGTGCATCTGCACAGGCAACGCGCGAGAGAGTGGAAAGTGGAGAGCA 537
Db 254 snLeuLeuGlnCysIleCysThrClyAsnGlyArgGlyGluTrpLysCysGluArGTrp 274
Qy 538 CTTCTGTGACAGACCAATGAGCGGATCTGGCCCTTCAACGATGTTGATTTGCT--- 593
Db 274 hrSerValGlnThrThrSerSerGlySerGlyProPheThrAspValArGAlaValT 294
Qy 593 ----- 593
Db 294 yrGlnProGlnProHisProGlnProProTyrcGlyHisCysValThrAspSerGly 314
Qy 594 -----GGACGTGAGTGGCTGTAGACCGTCCATCTGTCACAACAGCCCAT 639
Db 314 alValTyrcSerValGlyMetGlnTrpLeu-----LysThrGlnGlyAsnLysGlnM 331
Qy 640 TGGTTGTAGCTGTGCTGTAAGTGTGAGGCGGACGAATCAAGACATTAGTCTTAATTTT 699
Db 331 etLeuCysThrCysLeuGly----- 337
Qy 700 TTGAATTCATCTTAATCAATCAACCTGCTCATGAGAGAAAGACAGACAGCTTAAGTC 759
Db 338 -----AsnGlyValSer- 341
Qy 760 CAAATCAAAAACATTTGCTACTGATGATGCGCGGATGTCAATAACTTGAGAAACCTG 819
Db 341 ----- 341
Qy 820 ACTTAATAAGGCTATTCAAGAACATGATGCTGAACGTCAACAGTACAGACGACTACT 879
Db 342 -----CysGlnGluThrAlaValThrGlnThrTyrcGlyGlyAsnSerArg 357
Qy 880 TTGAG-----GTCAATTGATTTTGCAGGACGATGCAACCATTAATGATCGAAACGCGAAG 933
Db 357 lYgluProCysValLeuProPheThr-----TyrcAsnGlyArgT 370
Qy 934 TCTACTTTGCT-----GACAAAGATGTTGGTAACTTGGCCAGCCCAAC 978
Db 370 hrPheTyrcSerCysThrThrGlnGlyArGTrpAsnArgGlyHisLeuTrpCysSerThrHis 390
Qy 979 CTGTCAAGAAATTTTGTCTAAGCGGACATGCGCGTTAGACCATTAATAAGAAACCA 1038
Db 390 erAsnTyrcGlu----- 393
Qy 1039 TACAAAACCAAGGAAATCTGTGATGTGGAATATATCTGTACAGTTTATCCCTTAAC 1098
Db 394 --GlnAspGlnLysTyrcSerPheCysThrAspHieThrVal----- 406
Qy 1099 CTGATACGATTTCAACACAGTCTCAAAAGATCTAAGTATTGAAAAACATGACTATG 1158
Db 406 ----- 406
Qy 1159 GTGACACCATCATCTCAAGATTACTAGCTGACAGCAAAAGCATTTTAAACAAAAC 1218
Db 407 -----LeuValGlnThrArGly----- 412

```



```

QY 1219 ACCGAGGCTATAGATTATGAAAGCTGCTCATCGTCACTGATGACATGATTT 1278
DB 413 -----GlyAsnSerAsnGlyAlaLeuCysHis----- 421
QY 1279 TCCGTCAGATTTTACCAATGATTCAGAGCTTATCTACCGTGTAAATAATCGGAAACAG 1338
DB 422 -----PheProHeuLeuTyraAsnHis----- 429
QY 1339 CTTATAGATCAATAAAAATCTGCTGTGAATGAAGAAATTAACAACCTGACCTGATCT 1398
DB 430 -----AsnTyThrAspCysThrS----- 436
QY 1399 CTGAGAAATATTACGCTTAAATAAGGGAAAAAGCGTATGATCCCTTGTGATCGCAGTC 1458
DB 436 etGluGly-----ArgArgAsp 442
QY 1459 ACTGGA---CTGTTCACCATCAATATACGTTGATGTCGATCAAC---G 1503
DB 442 snMetLysTrpCysGlyThrGlnAsnTyraSerAlaSerGlnLysPheGlyPheCysAr 462
QY 1504 AATTGCTTAAAAAGTGAAGCTCTTAAACAGCTAGCGAAGTAACTTAAAGACTTCAAGATTT 1563
DB 462 rometalalaHisGluGluLeuCysThrThrAsnGlu----- 474
QY 1564 TATAGATCCGTGTATAGGTAAGGTAATCTATCAACAATCTCGATGCTTTGTGATTA 1623
DB 475 -----GlyVal 477
QY 1624 TGCACTATATCTTAACTGGAATAAGTAAAGATATATCAGATGACAAACCGTATATTA 1683
DB 477 et--TyraGlyLeuGlyAsnGlnTrpAspLysGlnHisAspMetGlyHisMetMetAlaGc 496
QY 1684 CCGTTTATATGGGCAAGCCGCAAGAGAGAAATGCTAGTCAATTAAGCTGTGTGTG 1743
DB 496 yethCysValGlyAsn-----GlyArg 504
QY 1744 GCCAGGCCCAACAGATTGA-----CCATAGCTGGAAGATGTTTGTATCATGCTGCTG 1797
DB 504 LysIleTrpThrCysIleAlaTySerGlnLeuArgSerGlnCysIle-----ValAsp 522
QY 1798 GACATCTCATGTGTCGGAAGAAAGTGGAGAAAGCCCTACCAAGGCTGGATGATGTGTG 1857
DB 522 spilletThrTyraAsnValAsnAspThrPheHisLysArgHisGluGluGlnHisMetLeu 542
QY 1858 ATTGACTTTCCTGCGAAGAGGCAAGGCAAGCATCACTTGCACTTGTAGAAATAGATGA 1917
DB 542 snCysThrCysPheGlyGlnGlyAlaGlyArgTrpLysCysAspProValAspGlnCysG 562
QY 1918 AGCATCAGACACAGAGACATCTATAGAAATGGAACACCTGGAAGCAAGAGATTAATC 1977
DB 562 LysAspSerGlnTrpThrPheTyrglnIleGlyAspSerTrpGluTyraValHis 581
QY 1978 GAGGAACCTGCTCCAGTGCATCTGACAGCAAGCAAGCGGAGAGAGTGAAGTGAAGA 2037
DB 582 --GlyValArgTyrglnCysTyrglyAlaGlyAlaGlyGlnIleGlyGlnTrpHisCysGln 600
QY 2038 GGCACACCTCTGTGACAGACACATGACGAGGATCTGGCCCC 2078
DB 601 -----ProLeuGlnThrTyroSerSerSerGlyPro 611

```

```

; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 627
; LENGTH: 2355
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-627

Alignment Scores:
Pred. No.: 2,48e-40 Length: 2355
Score: 634.50 Matches: 207
Percent Similarity: 36.69% Conservative: 70
Best Local Similarity: 27.42% Mismatches: 151
Query Match: 17.03% Indels: 328
DB: 6 Gaps: 28

US-09-940-235-12 (1-2096) x US-10-995-561-627 (1-2355)
QY 5 GACCATTCATGTTGTTGTCAGGTCGACAGCTTTTGCAGACAGAGTCGCTTACGCT--- 61
DB 122 AspCysThrCysIleGlyAlaGlyArgGlyAlaSerCysThrIleAlaAsnArgCys 141
QY 62 -----TCGCTCGGCTATCGGTGATTCATCTGCTTAAACAGTAAGCAACC 106
DB 142 HisGluGlyGlyGlnSerTyrgLysIleGlyAspThrTrp-----ArgArgPr 157
QY 107 CCGCCAGCTTACCGCGGCTCCTCAACAGACAGACAGATATGCGCACCCGTGGCCAGA 166
DB 157 ohLeuGluThr----- 160
QY 167 CCGAAGCTCCCGAGATCTCGATCCCGCGAATTAATAGACTCATATAGGAGACA 226
DB 160 ----- 160
QY 227 CAACGTTTCCCTTAGAA---ATAATTGTTTAATTAAAGAGATATATCAATGG 283
DB 161 -GlyGlyTyrgMetLeuGlnCysValCysLeuGlyAsnGlyLysGlyLys-TrpThr----- 178
QY 284 TGCAAGCAACAAGATTGTACCAATAGCTGGAAGATGTTTGTATCATGCTGCTGGAATT 343
DB 179 -----CysLysProIleAlaGlnLysCysPheAspHisAlaAlaGlyThrS 194
QY 344 CCTATGTCGTCGGAAGAAAGCTGGA----- 368
DB 194 etTyraValAlaGlyGluThrTrpGluLysProTyrglnGlyTyrmMetValAspCys 214
QY 369 -----GAAGCAGCGGACAGCATCACTTGCACTTGTAGAAATAGATGCAACGATC 417
DB 214 hrcysLeuGlyGluGlySerGlyArgIleThrCysThrSerAlaGlyAsnArgCysAsnArg 234
QY 418 AGGACACAGACATCTTATGAATTTGAGACACCTGAGACAAAGAGATATATGAGAAA 477
DB 234 LysPThrArgThrSerTyraGlyLeuLysPThrTrpSerLysLysAspAsnArgIya 254
QY 478 ACCTGCTCCAGTGCATCTGACAGGCAAGCGGAGAGAGTGAAGTGAAGGACACA 537
DB 254 snLeuLeuGlnCysIleCysThrGlyAsnGlyAlaGlyGlnTrpLysCysGlnArgHis 274
QY 538 CCTGTGACAGACACATGACAGGAGATCTGACCCCTTACACGATGTTGATTTGCT--- 593
DB 274 hserValGlnThrThrSerSerGlySerGlyProPheThrAspValAlaVal 294
QY 593 ----- 593
DB 294 tyGlnProGlnProHisProGlnProProProTyrglnHisCysValThrAspSerGly 314
QY 594 -----GACCTGAGTGGTCTGTAAACCTCATCTGTCAACAACGCAAT 639
DB 314 alValTyraSerValGlyMetGlnTrpLeu-----LysThrGlnGlyAsnLysGlnM 331
QY 640 TGGTGTATAGCGTTGCTGCTACTGTTGAGGGGACAGATCAAGACATTAATGCTTAATTTT 699
DB 331 etLeuCysThrCysLeuGly----- 337

```

```

QY 700 TTGAATCGATCTTAACATCAACGCTGCTATGAGGAGAAAGACAGACGAGCTTAAGTC 759
DB 338 -----AsnGlyValSer- 341
QY 760 CAAATCAAAACCATTTGCTACTGATAGTGGCCGATGTCAATAACTTGAGAAAGCTG 819
DB 341 ----- 341
QY 820 ACTTACTAAAGGCTATTCAAGAACAAATTGATCGTAAAGTCCACAGTAAGACGACTACT 879
DB 342 -----CysGlnGlnThrAlaValThrGlnThrTyrglyGlyAsnSerAsnG 357
QY 880 TTGAG-----GTCAATGATTTTGGCAAGCGATGCAACCATTAAGTATGCAAAACGCGAAG 933
DB 357 LylGlnProCysValLeuProPheThr-----TyraAsnIlyArgT 370
QY 934 TCTACTTTGCT-----GACAAAGATGGTTCGGTAACCTTGCCGACCCGAC 978
DB 370 hPheTySerCysThrThrGlnGlyArgGlnAspGlyIleLeuThrCysSerThrHis 390
QY 979 CTGTCCAAAGATTTTGGTAAAGGACATGTGGCGCTTAAGACCATATTAAGAAACCAA 1038
DB 390 eRaAsnTyrglu----- 393
QY 1039 TACAAACCCAGGAAATCTGTGATGTGAAATATATCTGACAGTTTACTCCCTTAACC 1098
DB 394 --GlnAspGlnIlySerTySerPheCysThrAspHisThrVal----- 406
QY 1099 CTGATGACGATTTGACAGCCAGGTCTCAAGATATAGTATTGAAAACTAGCTATCG 1158
DB 406 ----- 406
QY 1159 GTGACACCATCATCTCTCAAGAAATTACTAGCTCAAGACAAAGCATTTTAAACAAACC 1218
DB 407 -----LeuValGlnThrArgGly----- 412
QY 1219 ACCCAGGCTATACGATTTATGAAAGTGACCTCAATCGTCACTGACATGACATGACATTT 1278
DB 413 -----GlyAsnSerAsnGlyAlaLeuGlyHis----- 421
QY 1279 TCCGATGATTTTACCAATGATCAAGAGTTTACTTACCTGTGTAATAATCGGGAACAG 1338
DB 422 -----PheProPheLeuTyraAsnAsnHis----- 429
QY 1339 CTATATGATCAATAAAAATCTGCTGTGATGATGAAGAAATAAACAACACTGACTGATCT 1398
DB 430 -----AsnTyrgThrAspCysThrS 436
QY 1399 CTGAGAAATATTACGTCCTTAAAAAGGGGAAAGCCGTATGATCCCTTGTGATCGCAGTC 1458
DB 436 eGlnGly-----ArgAlaGAspA 442
QY 1459 ACTTGAAG--CTGTTCACCATCAATACGTTGATGTCATCAAC-----G 1503
DB 442 smMeLysrTyprCysGlyThrThrGlnAsnTyraAspAlaAspGlnIlyPheCysP 462
QY 1504 AATTGCTAAAAAGTGACAGCTCTTAAACAGCTAAGCAAGTAACTTAAGACTTGAGAGATT 1563
DB 462 rometAlaAlaHisGlnGlnIleCysThrThrAsnGln----- 474
QY 1564 TATAGATCTCTGTGATTAAGGCTAACTACTTACAAACAATCTCGATGCTTTTGGTATTA 1623
DB 475 -----GlyValM 477
QY 1634 TGGCATATCTTAACTGAAGAAAGTAGAGATATACAGATGACACCAACCGTATGATA 1683
DB 477 eTyrArgIleGlyAspGlnTTPAspGlnHisAspMetGlyHisMetMetArgC 496
QY 1684 CCGTTATATGGGCAAGCAGACCCGGAAGAGAGAAATCTAGCTACATTTAGCTGTGTGTG 1743
DB 496 ySThrCysValGlyAsn-----GlyArgS 504

```

```

QY 1744 GCCAGGCGCAACAGATTGTA-----CCCATAGCTAGAGAGTGTTTGATCATGCTGCTG 1797
DB 504 LylGlnThrThrCysAlaLeaIleTySerGlnLeuAspGlnCysIle-----ValAspA 522
QY 1798 GCACTTCCTATGTCGTGGAGAAACGTGGAGAAACCTTACCMAAGCTGTGATGTAG 1857
DB 522 spIleThrTyraValAsnAspThrPheHisIlyArgHisGlnGlnIlyHisMetLeuA 542
QY 1858 ATTGACTTGGCTGGGAGAGGAGGAGCCATCACTTGCACTTGAATAGATGCA 1917
DB 542 smCysThrCysPheGlyGlnIleValArgIlyArgIlyPheCysAspProValAspGlnCysG 562
QY 1918 ACATTCAGACACAAAGACATCTTATGAATTTGAGACACCTTGAGACAAAGATATTC 1977
DB 562 InAspSerGlnThrGlyThrPheTyrglnIleGlyAspSerThrGlnIlyTyraHis- 581
QY 1978 GAGGAAACCTGCTCCAGTCATCTGCAAGGCAACGCGGAGAGAGATGGAAGTGTGAGA 2037
DB 582 --GlyValArgTyGlnCysTyrcCysTyrgIlyArgGlyIleGlyGlnThrHisCysGln- 600
QY 2038 GGCACACTCTGTGACAGACCATCGAGCGGATCTGGGCCC 2078
DB 601 -----ProLeuGlnThrTyrgProSerSerSerGlyPro 611

RESULT 8
US-10-821-234-1545
; Sequence 1545, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; PRIOR FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pc_seq_genes Version 1.0
; SEQ ID NO 1545
; LENGTH: 2384
; TYPE: PRF
; ORGANISM: Homo sapiens
; US-10-821-234-1545

Alignment Scores:
Pred. No.: 2,48e-40 Length: 2384
Score: 634.50 Matches: 207
Percent Similarity: 36.69% Conservative: 70
Best Local Similarity: 27.42% Mismatches: 151
Query Match: 17.03% Indels: 328
DB: 6 Gaps: 28

US-09-940-235-12 (1-2096) x US-10-821-234-1545 (1-2384)
QY 5 GACCATTCATGTTGTTGCTCAGATCGACAGCTTTGACAGACAGTCTTGACGT--- 61
DB 151 AspCysThrCysIleGlyAlaGlyArgGlyArgIleSerCysThrIleAlaAsnArgCys 170
QY 62 -----TCGCTCGCGTATCGGTGATTCATTCTGTAACCAAGTAAGCAACC 106
DB 171 HisGlnGlyGlyGlnSerTyrgLysIleGlyAspThrTrp-----ArgArgPr 186
QY 107 CCGCCAGCCTAAGCCGGGCTCTCAACGACAGAGACATATGCGCACCCGTGCGCAGGA 166
DB 186 ohIleGlnThr----- 189
QY 167 CCGAAGCTCCCGAGATCTCGATCCCGGAATTAAAGATCACTATAGGAGACGA 226
DB 189 ----- 189

```



```

Db      496 ysthrCyValglYasn-----GlyArg 504
Qy      1744 GCCAGCGCGCAAGATTGTA-----CCATAGCTGAGAGCTTTTGTATCATGCTGCG 1797
Db      504 lylglutrrpThrCysllealalyrSerglnleuargspglnCysile-----ValAspA 522
Qy      1798 GCATCTCTATGTCGTGCGAGAAACGTGGAGAACCCCTACCAAGCTGGATGATGATG 1857
Db      522 splleThrrTyrrsValaIaenAsprrrPhehileylarghileglugllyhileMeleuA 542
Qy      1858 ATTGTACTTCTGCTGGAGAGACGCGACGCGATCATCTTGCATCTTGTAGAAATAGTGA 1917
Db      542 snCystrhCySrhelgylngllyArgglYArgTrpIysCyAsprrProValAspGlnCySg 562
Qy      1918 ACATCAGACACACAGACATCTTATGAAATGGACACCTGGACCAAGAGATTAATC 1977
Db      562 lnaAspSerglnThrrlyThrrprrlyrGlnlleglYAspserTrpIulysrYValhls- 581
Qy      1978 GAGCAACCTGCTCGATGTCATCTGACAGGACGCGGAGAGAGAGTGTGAGA 2037
Db      582 --GlyValArgTrpGlnCystrYrGlyArgglYlleglYglutrrPhisCySgln- 600
Qy      2038 GGCAACCTCTGTGACACCATGCGAGCGGATCTGCCCC 2078
Db      601 -----ProleuglnThrrYrProserSerserClyPro 611

```

```

RESULT 10
US-10-631-558-4
; Sequence 4, Application US/10631558
; Publication No. US20050260598A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammarra
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/10/631,558
; CURRENT FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US/09/940,235
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-631-558-4

```

```

Alignment Scores:
Pred. No.: 2,5e-40 Length: 259
Score: 632.00 Matches: 117
Percent Similarity: 87.59% Conservative: 3
Best Local Similarity: 85.40% Mismatches: 11
Query Match: 16.96% Indels: 6
DB: Gaps: 3

```

```

US-09-940-235-12 (1-2096) x US-10-631-558-4 (1-259)
Qy      1698 AAGCGACCCGAGAGAGATGCTAGTACCATTTA-----GCTGGTGTGCG 1745
Db      124 ArgArgProHis---GlnThrrGlnGlyTrpMetleuGlnCyValCysleuGlyasnGly 142

```

```

Qy      1746 CAGCGCGAA---CAGATTGTACCATTGATGAGAGGTTTGTATCATGCTGAGACT 1802
Db      143 lylglYglutrrpThrCyslySprollealGlnlyCySrhAsprrlsalalaglYThr 162
Qy      1803 TCCTATGTGTGTCGAGAGAAAGTGGAGAACCCCTACCAAGCTGGATGATGATG 1862
Db      163 SerTyValValGlyglutrrTrpIulysProTrpGlnGlyTrpMetMetValAspCys 182
Qy      1863 ACTTGGCTGGAGAGACGCGACGCGATCATCTTGCATCTTGTAGAAATAGTGAACGAT 1922
Db      183 ThCySleuGlnGlygluGlySerglyArglleThrCySrhSerrArgAsnArgCySAsnAsp 202
Qy      1923 CAGCAACAAGACATCTTATGAAATGGAGACACCTGGACCAAGAGTAATCGAGA 1982
Db      203 GlnAspThrrArgThrrSerTyrrArglleGlyAsprrThrrprrlySlyAspAsnArgGly 222
Qy      1983 AACCTCTCCAGTGCATCTTGCACAGGACGCGCGAGAGAGTGAAGTGAAGAGCAC 2042
Db      223 AsnleuGlnCyslleCysThrrGlyAsnGlyArgglYglutrrpIysCySgluArgHis 242
Qy      2043 ACCTTGTGACACCATGTCAGCGGATCTGGCCCTTACCGATGTTGCT 2093
Db      243 ThrrSerValGlnThrrThrrSerSerglySerglyProPheThrrAspValArg 259

```

```

RESULT 11
US-11-006-119-31
; Sequence 31, Application US/11006119
; Publication No. US20050260691A1
; GENERAL INFORMATION:
; APPLICANT: Ndao, Momar
; APPLICANT: Ward, Brian
; APPLICANT: Caffrey, Rebecca
; APPLICANT: Spithill, Terry
; APPLICANT: Li, Hongshan
; APPLICANT: Podust, Vladimir
; APPLICANT: Perichon, Regis
; APPLICANT: CIPHERGEN Biosystems, Inc.
; TITLE OF INVENTION: Serum Biomarkers for Chagas Disease
; FILE REFERENCE: 016866-012130US
; CURRENT APPLICATION NUMBER: US/11/006,119
; CURRENT FILING DATE: 2004-12-06
; PRIOR APPLICATION NUMBER: US 60/527,153
; PRIOR FILING DATE: 2003-12-05
; PRIOR APPLICATION NUMBER: US 60/565,093
; PRIOR FILING DATE: 2004-04-22
; PRIOR APPLICATION NUMBER: US 60/625,519
; PRIOR FILING DATE: 2004-11-06
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 31
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: N-terminal 28.7 fragment of Fibronectin
; US-11-006-119-31

```

```

Alignment Scores:
Pred. No.: 2,57e-40 Length: 300
Score: 632.00 Matches: 117
Percent Similarity: 87.59% Conservative: 3
Best Local Similarity: 85.40% Mismatches: 11
Query Match: 16.96% Indels: 6
DB: Gaps: 3

```

```

US-09-940-235-12 (1-2096) x US-11-006-119-31 (1-300)
Qy      1698 AAGCGACCCGAGAGAGATGCTAGTACCATTTA-----GCTGGTGTGCG 1745
Db      124 ArgArgProHis---GlnThrrGlnGlyTrpMetleuGlnCyValCysleuGlyasnGly 142
Qy      1746 CAGCGCGAA---CAGATTGTACCATTGATGAGAGGTTTGTATCATGCTGCGGACT 1802

```

```

Db      143  LysGlyGluThrCysLysProIleAlaGluLysCysPheAspHisAlaAlaGlyThr 162
Qy      1803  TCTATGATGTCGAGAGAAAGTGGAGAGAGCCCTACCAAGGCTGATGATGATGATGAT 1862
Db      163  SerTyValValGlyGluThrTrpGluLysProIleGlnGlyTrpMetCysValAspCys 182
Qy      1863  ACTTCCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1922
Db      183  ThrCysLeuGlyGluGlySerGlyArgGlyLeuThrCysThrSerArgAsnAlaGlyAsnAsp 202
Qy      1923  CAGGACACAAAGACATCTTATGAATTTGAGACACCTCGAGACAGAAAGATATCGAGA 1982
Db      203  GlnAspThrArgThrSerTyArgGlyLeuAspThrTrpSerTyLysAspAsnArgGly 222
Qy      1983  AACCTGCTCAGGACATCTGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2042
Db      223  AsnLeuLeuGlnCysIleCysTrnGlyAsnGlyArgGlyGluTrpLysCysGluArgHis 242
Qy      2043  ACCTGTGTCAGACACATCGAGAGAGATGAGCCCTTCACCGATGTTGCT 2093
Db      243  ThrSerValGlnThrThrSerSerGlyProPheThrAspValArg 259

RESULT 12
US-10-995-561-631
; Sequence 631, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 631
; LENGTH: 642
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-995-561-631

Alignment Scores:
Pred. No.: 2,97e-40 Length: 642
Score: 632.00 Matches: 117
Percent Similarity: 87.59% Conservative: 3
Best Local Similarity: 85.40% Mismatches: 11
Query Match: 16.96% Indels: 6
DB: Gaps: 3

US-09-940-235-12 (1-2096) x US-10-995-561-631 (1-642)
Qy      1698  AAGCGACCCGAGAGAGATGCTAGCTACATTTA-----GCTGGTGTGTC 1745
Db      155  ArgGluProHis-----GluThrGlyGlyTyMetLeuGlnCysValCysLeuGlyAsnGly 173
Qy      1746  CAGGCGCAA--CAGATTGTACCATAGCTAGAGAGTGTATGATCACTGCTGGAGCT 1802
Db      174  LysGlyGluThrTrpCysLysProIleAlaGluLysCysPheAspHisAlaAlaGlyThr 193
Qy      1803  TCTATGATGTCGAGAGAAAGTGGAGAGAGCCCTACCAAGGCTGATGATGATGATGAT 1862
Db      194  SerTyValValGlyGluThrTrpGluLysProIleGlnGlyTrpMetCysValAspCys 213
Qy      1863  ACTTCCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1922
Db      214  ThrCysLeuGlyGluGlySerGlyArgGlyLeuThrCysThrSerArgAsnAlaGlyAsnAsp 233
Qy      1923  CAGGACACAAAGACATCTTATGAATTTGAGACACCTCGAGACAGAAAGATATCGAGA 1982
Db      234  GlnAspThrArgThrSerTyArgGlyLeuAspThrTrpSerTyLysAspAsnArgGly 253
Qy      1983  AACCTGCTCAGTGCATCTGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2042

```

```

Db      254  AsnLeuLeuGlnCysIleCysThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHis 273
Qy      2043  ACCTGTGTCAGACACATCGAGAGAGATGAGCCCTTCACCGATGTTGCT 2093
Db      274  ThrSerValGlnThrThrSerSerGlyProPheThrAspValArg 290

RESULT 13
US-10-995-561-621
; Sequence 621, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 621
; LENGTH: 1341
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-995-561-621

Alignment Scores:
Pred. No.: 1.02e-13 Length: 1341
Score: 278.50 Matches: 136
Percent Similarity: 34.32% Conservative: 73
Best Local Similarity: 22.33% Mismatches: 195
Query Match: 7.47% Indels: 205
DB: Gaps: 25

US-09-940-235-12 (1-2096) x US-10-995-561-621 (1-1341)
Qy      397  CTAGAATAATAGATGACAGATCAGACACAAAGACATCTTATGAATTTGAGACACCTGGA 456
Db      784  MetGluSerGlnProLeuIleGlyThrGlnSerThrAlaIle-----ProAlaProThr 801
Qy      457  GCAAGAGATATATCGAGAGAAAGCTGCTCCAGTGCATCTGACAGAGAGAGAGAGAGAG 516
Db      802  AsnLeuLysPheThrGlnValThrProThrSerLeuSerAlaGlnTrpThr-ProProAs 821
Qy      517  AGTGAAGTGTGAGGAGACACCTCTGTGACACACATCGAGAGAGAGATCTGAGCCCTTCA 576
Db      821  nValGlnLeu-ThrGlyTyArgValArgValThrProLysGluLysThrGlyProMetL 841
Qy      577  CCGATGTTGCTATGCTGAGACCTGAGTGGCTGCTAAGCCGTCATCTGTCAACAAGCC 636
Db      841  LysGluLeuAsnLeuVala-----ProAspSerSerS 851
Qy      637  AATTGCTTTAGGCTGTGCTAGCTGTGAGGAGAGAGATCAGACATTTAGCTTAAT 696
Db      851  erValValValSer-----GlyLeuMetValAlaThrLysTyGlnValSerValTyra 869
Qy      697  TTTTGAATTCATTAACATCAGACCTGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 753
Db      869  IaleuLysAspThrLeuThrSerArgProAlaGlnGlyValValThrThrLeuGlnLysn 889
Qy      754  TAACTCCAAATCAAAAACATTTGCTATCTGATAGTGGAGAGAGATGACATTAACCTTGA 813
Db      889  alSerProProAlaArgAlaArgValThrAsp----- 899
Qy      814  AAGCTGACTTACTTAAAGCTATTCAGAAACATTTGATCGCTAAGTCCACAGTAACGAG 873
Db      899  ----- 899
Qy      874  ACTACTTGAAGTCAATTTTGAAGAGATGACACATTAATCTGATGAAACGAGAGAG 933
Db      900  -----AlaThrGluThrThrIleThrIle----- 907

```

```

Qy 934 TCTACTTGTGTCGAAAGATGGTTCGGTAACCTTGCAGCCAGCAACCTGTCCAGAAATTT 993
Db 908 -----SerTrpArgThrIleuThrIleThrglyPheg 920
Qy 994 TGCCTAACGCGACATGCGCGCTTATGACCATATTAAGAAAAACCATTAACAAACCAAGCA 1053
Db 920 InValAspAla-----ValProAlaAsnGlyInThrProIleGlnArgThrIle 937
Qy 1054 AATCGTGTGATGGAAATATGCTACGATTACTCCCTTAAACCTGATGAGATTTCA 1113
Db 937 YSPProAspValArgSerTrpThrIle-----ThrIleuGlnProGlyThrAspTrp 955
Qy 1114 GACCAAGTCTCAAGATTAAGTAACTATTTGAAAACTAGCTATCGGTGACACCATCATCAT 1173
Db 955 YS-----IleTrpLeuTrpThrIleuAsnAspAsnAlaArgS 967
Qy 1174 CTCAGAAATTAAGTCTCAAGCA----- 1196
Db 967 erSerProValValIleAspAlaSerThrAlaIleAspAlaProSerAsnLeuArgPheL 987
Qy 1197 -----CAAGCATTTTAAAGAAAAACCAACCA-----G 1224
Db 987 euAlaIleThrProAsnSerLeuLeuValSerTrpGlnProProArgAlaArgIleThrg 1007
Qy 1225 GCTATACGATTT--TATGAAGCTGACCTCTCAATCGTCACTCATGACATGACATTTTCC 1281
Db 1007 IYTrIleIleIleuSerTrpGlnuSerProGlySerProPro-----A 1020
Qy 1282 GTACGATTTTCAACATGATGCA----- 1304
Db 1020 rGlnuValValProArgProArgProGlyValThrGlnuAlaThrIleThrgIleuGlnu 1040
Qy 1305 -----GAGTTTACTTACCGTGT-----AAAAATGGGAGCAAGCTTATAGCA 1347
Db 1040 rGlnuIleuThrIleuTrpIleuTrpIleuValIleAlaLeuAlaSerAsnGlnuSerGlnuProL 1060
Qy 1348 TC-----AATAAAAAATCGTGTGTAATGAAGAAATTAACACACGACCTGATCTCG 1401
Db 1060 euIleGlyArguGlySerThrGly--GlnuGlnuAlaSerGlnuThrIle----- 1076
Qy 1402 AGAAATATTAAGTCTTAAAAAAGGGGAAAGCCGATGATCCCTTGTATCGCAGTCACT 1461
Db 1077 -----SerTrpAlaProPheGln----- 1082
Qy 1462 TGAACCTGTTCACCATCAATACGTTGATGTGATTCGACCAAGATTCCTTAAAAAGTACG 1521
Db 1083 -----AspThrSerGlnuTrpIleIleSerCysH 1092
Qy 1522 AGCTCTTAACAGCTAGCGAAAGTAACTTAGACTTCGAGATTTATTCAGATCCGCGATA 1581
Db 1092 IsProValGlyThrAspGlnuGlnuProLeuGlnuPheArg-----ValProGlyThrs 1109
Qy 1582 AGGCTAACTACTTACAAACAATCTCGATGCTTTGGTATTATGACTATTAACCTTAACTG 1641
Db 1109 erThrSerAla-----ThrLeuThrd 1116
Qy 1642 GA-----AAAGTAGAGCAATATC 1659
Db 1116 IYLeuThrArgGlyAlaThrTrpAsnIleIleValGlnuAlaLeuIYAspGlnuArgH 1136
Qy 1660 AGGATGACCAACCGATCATACCGTTATATATGCGCAAGGACCCGAGAGAGAGAAAG 1719
Db 1136 IsIYAspValArgGlnuGlnuValIleValIleValGlyAsnSerValAsn--GlnuGlyLeuAsnG 1155
Qy 1720 CTAGCTACCATTTAGCTGTGTGTGCGCAGCGCAACAGATTGATCCCATAGCTGAGAAAGT 1779
Db 1155 In-----ProThrAspAspSerC 1161
Qy 1780 GTTTGATCATCTGCTGGGACTTCTATGTGTGCGAGAAACGTGGAGAGAGCCCTTACC 1839
Db 1161 YSPProAspProTrpThrValSerThrIleValGlyAspGlnuTrpGlnuArgMetSerC 1181
Qy 1840 AAGCGTGATGATGTAGATTCTTACTTGTGCTGGAGAGAGCGAGCCGACCATCATCTTGA 1899

```

```

Db 1181 IuSerGlyPheIYLeuLeuCysGlnuIYLeuGlyPheGlySerGlyHIsPheArgCysA 1201
Qy 1900 CTCTTAGAAATTAAGTCAACGATCAAGACACAAAGACATCTTATGAATTTGAGACACT 1959
Db 1201 erSerArgTrpCysHIsAspAsnGly-----ValAsnTrpIleGlyGlnuYst 1219
Qy 1960 GGAGCAAGAGATTAATTCAGAGAAACCTGCTCCAGTGCATCTGCACAGGCAAGCGCGAG 2019
Db 1219 rPAspArgGlnuGlyGlnuAsnGlyGlnuMetSerCysThrCysLeuGlyAsnIYIYusG 1239
Qy 2020 GAGAGTGAAGTGTGAGAGGAC 2042
Db 1239 IYGlnuPheIYCysAspProHIs 1246

RESULT 14
US-10-995-561-632
; Sequence 632, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: C1001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 632
; LENGTH: 693
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-995-561-632

Alignment Scores:
Pred. No.: 1,51e-13 Length: 693
Score: 275.50 Matches: 142
Percent Similarity: 32.88% Conservative: 77
Best Local Similarity: 21.32% Mismatches: 220
Query Match: 7.39% Indels: 227
DB: Gaps: 25

US-09-940-235-12 (1-2096) x US-10-995-561-632 (1-693)
Qy 346 TATGTGTCGAGAAACGTGGAGAAAGCAGCGACGATCACTTGACCTTATGAATA 405
Db 61 TrValValSerValTrpAlaGlnAsnProSerGlyIuSerGlnProLeuValGlnThr 80
Qy 406 GATGCAACGATCAGACACAAAGACATCTTATGAATTTGAGACACTGAGCAAGAG 465
Db 81 AlaValThr-----ThrIleProAlaProThrAspLeuIYs 92
Qy 466 ATATGAGGAACCTGCTCCAGTGCATCTGCACAGGACCGCGAGAGTGAAGT 525
Db 93 PheThrGlnuValThrProThrSerLeuSerAlaGlnuThrProProAsnValGlnu 112
Qy 526 GTGAGAGGACACACCTGTGTCAGACCAATCGAGCGGATGCGCCCTTACCGAGTTC 585
Db 112 u-ThrglyTrpArgValArgValThrProIYSerGlnuSerThrGlyPrometIYSerIleA 132
Qy 586 GTATTGCTGACCTGAGTGCTGTAGACCGTCAATCTGTCAACAACAGCAATGTGTTG 645
Db 132 snLeuAla-----ProAspSerSerValValI 142
Qy 646 TTAGCGTGTGCTGACTGTGAGGGAGCAATCAAGACATTAGCTTAAATTTTGA 705
Db 142 alSer-----GlyLeuMetValAlaThrIYSerGlnuValSerValTrpAlaLeuIYsa 160
Qy 706 TCGATTAACATCAACCTGCTCATGAGAGAAAGACAGAG-----CAAGCTTAAGTCAA 762
Db 160 spThrIleuThrSerArgProAlaGlnuGlyValValThrThrIleuGlnuAsnValSerPro 180

```

```

QY 763 AATCAAAACCAATTTGCTAGTAGTGGCGCGATGTGCATTAACCTTGAGAAAGCTGACT 822
Db      |||
180 roArgArgValaArgValaThrAsp----- 187
QY 823 TACTTAAAGCTATTCAGAAACAATTGATGCGTAAACGTCCACAGTAACGAGACTTCTTGG 882
Db      ----- 187
QY 883 AGGTCAATTGATTTCGACAGATGCAACCTTACTGATCGAAACGGCAAGGTCTACTTGG 942
Db      |||
188 -----AlaThrGluThrThrIleThrIle----- 195
QY 943 CTGACAAAGATGGTTCGGTAACCTTCCGCAACCTGTCCAGAAATTTTGCTAAGC 1002
Db      |||
196 -----SerTrpArgThrIleThrGluThrIleThrGluPheGlnValaAsp 211
QY 1003 GACATGTGGCGCTTACACCATATTAAGAAACCAATACAAACCAAGCAATCTGTTG 1062
Db      |||
211 la-----ValProAlaAsnGlyGlnThrProIleGlnArgThrIleLeuProAsp 228
QY 1063 ATGTGAATATATCTGTACAGTTTACTCCCTTAACCCCTGATGACGATTTCAGACCAAGTCC 1122
Db      |||
228 laArgSerTrpThrIle-----ThrGlyLeuGlnProGlyThrAspTrpLys----- 243
QY 1123 TCAAGATATCTAAGCTATTGAAACCACTAGCTATCGGTGACACCATCATCTCAGAAAT 1182
Db      |||
244 -----IleTrpLeuTrpThrIleuAsnAspAsnAlaArgSerSerPro 258
QY 1183 TACTAGCTCAAGCA----- 1196
Db      |||
258 alValIleAspAlaSerThrAlaIleAspAlaProSerAsnLeuArgPheLeuAlaThr 278
QY 1197 -----CAAGCAATTTTAAACAAACCAACCA-----GGCTATATCA 1233
Db      |||
278 hrProAsnSerLeuLeuValSerTrpGlnProProAlaArgAlaIleThrGlyLeuIle 288
QY 1234 TT-----TATGAACGTGACCTCTCAATGCTCATGATGACATGACATTTTCCGTACGATTT 1290
Db      |||
298 leLeuTrpGluArgProGlySerProPro-----ArgGluVal 311
QY 1291 TACCAATGATCA-----G 1305
Db      |||
311 alProArgProArgProGlyValThrGluAlaThrIleThrGlyLeuGluProGlyThrG 331
QY 1306 AGTTTACTTACCGGTGTTAAATCGGAAACAACCTTATAGATCAATATAAAATCT----- 1361
Db      |||
331 luTrpThrIleTrpVal-----IleAlaLeuValAsnAlaAsnGlnLysSerGlu 347
QY 1362 -----GGTCTGAATGAAGAAATATAACAACACCTGACCTGATCTCGAATAATATTACG 1413
Db      |||
347 roLeuIleGlyArgGlyLysThrValGlnLysThrProPheValThrHisProGlyTrp 367
QY 1414 TCCTTAAAAAAGG----- 1427
Db      |||
367 spThrGlyAsnGlyIleGlnLeuProGlyThrSerGlyGlnInProSerValGlyGln 387
QY 1427 ----- 1427
Db      ----- 1427
QY 387 ImMeIlePheGluGluHisGlyPheArgArgThrThrProProThrAlaThrProI 407
QY 1428 -----GAAAGCCGTATGATCCC-----TTTATGCGCACTGCACTTGAACCTGTCA 1473
Db      |||
407 leArgHisArgProArgProGlyTrpProAsnValGlyGlnGluAlaLeuSerGlnThr 427
QY 1474 CCATCAAAATACGTTGATGTC-----GATACCAACGAATGTTAAAGAGAGACAGCTCTTAA 1530
Db      |||
427 hrIleSerTrpAlaProPheGlnAspThrSerGlyTrpIleIleSerCysHisProValG 447
QY 1531 CAGTAGCGAAGCAAGTAACTTGAATTCAGATTATATACATCTCGTATTAAGGCTTAAC 1590
Db      |||
447 lyThrAspGluGluProLeuGlnPheArg-----ValProGlyThrSerThrSer 464
QY 1591 TACTCTAACAAACATCTCGATGCTTTGGATTATATGACATATACCTTAACCTGA----- 1643

```

```

Db      |||
464 la-----ThrLeuThrGlyLeuThr 471
QY 1644 -----AAAGTAGAGATTAATCAACATGACA 1668
Db      |||
471 rgGlyAlaThrTyraAsnIleIleValGluAlaLeuLysAspGlnGlnArgHisLysVal 491
QY 1669 CCAACCGTATCATTAACGTTTATATGAGCAACGCAACCGAAGAGAGAAATGCTAGTACC 1728
Db      |||
491 rgGluGluValAlaThrValGlyAsnSerValAsn-----GluGlyLeuAsnGln----- 507
QY 1729 ATTTAGCTGTGTGGCCAGCGCAACAGATTTGATCCCATAGCTGAGAAAGTGTTCATC 1788
Db      |||
508 -----ProThrAspArgSerCysPheAsp 516
QY 1789 ATGCTGTGGGACTTCTATGTGTGCGGAGAAAGTGGAGAAAGCCCTTCAAGGCTGGA 1848
Db      |||
516 roTyThrValSerHisTyraIleValAlaGlyAspGluTrpGluArgMetSerGluSerGly 536
QY 1849 TGATGTGATGATTGACTTTCCTGAGAGAGAGCGGACGCGATCACTTGCACTTCTAGAA 1908
Db      |||
536 heLysLeuLeuCysGlnCysLeuGlyPheGlySerGlyHisPheArgCysAspSerSer 556
QY 1909 ATAGATCAACGATCAGACACAAAGACATCTTATAGATTGAGACACCTTGACAGAA 1968
Db      |||
556 rgTrpCysHisAspAsnGly-----ValAsnTrpLysIleGlyLysTrpAspArg 574
QY 1969 AGGATTAATCAGAGAAACCTGCTCCAGTGCATCTGCACAGGCAAGCGCCGAGAGATGGA 2028
Db      |||
574 lngLysGluAsnGlyGlnMetMetSerCysThrCysLeuGlyAsnGlyLysGlyGluPhe 594
QY 2029 AGTGTGAGAGGCAC 2042
Db      |||
594 yscYAspProHis 598

RESULT 15
US-10-995-561-630
; Sequence 630, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCES: C1001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 630
; LENGTH: 1315
; TYPE: PRN
; ORGANISM: Homo sapiens
US-10-995-561-630

Alignment Scores:
Pred. No.: 1,71e-13 Length: 1315
Score: 275.50 Matches: 142
Percent Similarity: 32.88% Conservative: 77
Best Local Similarity: 21.32% Mismatches: 220
Query Match: 7.39% Indels: 227
DB: Gaps: 25

US-09-940-235-12 (1-2096) x US-10-995-561-630 (1-1315)
QY 346 TATGTGTGCGAGAAAGTGGAGAAAGCGAGCCATTCATTCAGATAATA 405
Db      |||
683 TyraValAsnValTyraAlaGlnAsnProSerGlyLysArgInProLeuValGlnThr 702
QY 406 GATGCAAGATCAGACACAAAGACATCTATAGATTGAGACACCTGAGCAAGAAAG 465
Db      |||
703 AlaValThr-----ThrIleProAlaProThrAspLeuLys 714

```



```

Qy 466 ATAAATGAGAAACCTGCTCCAGTGCATCTGCACAGGCAACGCGCCGAGAGTGAAGT 525
Db      |||
Qy 725 PheThrGlnValThrProThrSerLeuSerAlaGlnThrProProAsnAlaGlnLeu 734
Db      |||
Qy 526 GTGAGAGGCAACCTCTGTGCAGACCAATCGACGGATGCGCCCTTCACCGATGTC 585
Db      |||
Qy 734 u-ThrGlyTyrArgValArgValThrProLeuGlyThrGlyProMetLeuGlnLeu 754
Db      |||
Qy 586 GTATTGTGAGACCTGAGTGGCTGTAGACCGTCCATCTGCACACCAAGCAATGTTG 645
Db      |||
Qy 754 snLeuAla-----ProAspSerSerValVal 764
Db      |||
Qy 646 TTAGCGTTCGTGTACTGTGTGAGGGAGCAATCAAGACATTAAGTCTTAAATTTTGA 705
Db      |||
Qy 764 alSer-----GlyLeuMetValAlaThrLysTyrGlnValSerValTyrAlaLeuLys 782
Db      |||
Qy 706 TCGATCTAACATCAGCACTGCTGCATGAGAGAAAGACAGAG--CAAGGCTTAAGTCA 762
Db      |||
Qy 782 spThrLeuThrSerArgProAlaGlnGlyValValThrThrLeuGlnAsnValSerPro 802
Db      |||
Qy 763 AATCAAAACCATTTGCTACTGATAGTGGCGCATGTCACATTAACCTTGAGAAAGCTG 822
Db      |||
Qy 802 roArgArgAlaArgValThrAsp----- 809
Db      |||
Qy 823 TACTAAAGGCTATTTCAGAAACAATTGATGCTAACGTCACAGTAAAGACGACTTGTG 882
Db      |||
Qy 809 ----- 809
Qy 883 AGGTGATTGATTTCGACAGCATGCAACCATTACTGATGCAACGCAAGTCTACTTGTG 942
Db      |||
Qy 810 -----AlaThrGlnThrThrLethrLethrLethrLethrLethrLethrLethr 817
Qy 943 CTGACAAAGATGTTGCTGTAACCTGCGCACCCACCACTGCTCCAGAAATTTTCTAAGC 1002
Db      |||
Qy 818 -----SerTyrArgThrLysThrGlnThrLethrGlnThrGlnValAlaAsp 833
Qy 1003 GACATGCGCGGTGTAGACCATATAAAGAAACCAATACAAACCAAGCGAATCTGTTG 1062
Db      |||
Qy 833 la-----ValProAlaAsnGlyGlnThrProLleGlnAlaGlnThrLethrProAsp 850
Qy 1063 ATGTGGAATATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTGACAGCAG 1122
Db      |||
Qy 850 alaGlySerTyrThrLethrLethrLethrLethrLethrLethrLethrLethrLethr 865
Qy 1123 TCAAAAGATACCTAGCTATTGAAACCACTAGCTATGCTGTCACACCATCACTTCAGAA 1182
Db      |||
Qy 866 -----LethrLethrLethrLethrLethrLethrLethrLethrLethrLethr 880
Qy 1183 TACTAGCTCAAGCA----- 1196
Db      |||
Qy 880 alValIleAspAlaSerThrAlaIleAspAlaProSerAsnLeuArgPheLeuAlaThr 900
Qy 1197 -----CAAGCATTTTAAACCAAAACCAAGCA-----GGCTATACGA 1233
Db      |||
Qy 900 hrProAsnSerLeuLeuValSerTyrGlnProProArgAlaArgIleThrGlyTyrLethr 920
Qy 1234 TT--TATGACGTGACCTCTCAATCGTACATGCATGACATGACATGACATTTCCGTA 1290
Db      |||
Qy 920 leLysTyrGlnLysProGlySerPro-----ArgGlnValVal 933
Qy 1291 TACCAATGATGACA-----G 1305
Db      |||
Qy 933 alProArgProArgProGlyValThrGlnAlaThrLethrGlyLeuGlnProGlyThrG 953
Qy 1306 AGTTTACTTACCGTGTAAATAATCGGAGAACAGCTTATAGATCAATTAATAATCT-- 1361
Db      |||
Qy 953 lutyThrLethrLethrVal-----IleAlaLeuLysAsnAlaGlnLysSerGln 969
Qy 1362 -----GTTCTGAATGAAGAAATTAACACACATGACCTGATCTCTGAGAAATATTA 1413
Db      |||
Qy 969 roLeuIleGlyArgLysLysThrValGlnLysThrProPheValThrHisProGlyTyr 989
Qy 1414 TCCTTAAAAAAGG----- 1427

```

```

Db      |||
Qy 989 spThrGlnAsnGlyIleGlnLeuProGlyThrSerGlyGlnGlnProSerValGlyGln 1009
Qy 1427 ----- 1427
Db      |||
Qy 1009 lmetIlePheGlnLysIleGlyPheArgArgThrThrProProThrThrAlaThrPro 1029
Db      |||
Qy 1428 -----GAAAGCCGATGATCCC--TTGATCGCAGTCACTTGAACGTTCA 1473
Db      |||
Qy 1029 leArgHisArgProArgProTyrProProAsnValGlyGlnGlnAlaLeuSerGlnThr 1049
Qy 1474 CCATCAATACGTTGAAGTGC--GATACCAACGAATTCCTAATAAGTACGAGCTTTAA 1530
Db      |||
Qy 1049 hrLieserThrAlaProPheGlnAspThrSerGlyTyrLethrLethrLethrLethr 1069
Qy 1531 CAGCTACGAAAGCTTAACCTTACACTTCAAGATTTATACATCTCTGATTAAGCTTAA 1590
Db      |||
Qy 1069 lYThrAspGlnGlnProLeuGlnPheArg-----ValProGlyThrSerThrSera 1086
Qy 1591 TACTCTACAAACATCTGATGCTTTGATATATAGACTATACCTTAACCTGGA----- 1643
Db      |||
Qy 1086 la-----ThrLeuThrGlyLeuThrAla 1093
Qy 1644 -----AAAGTAGAGATATACAGCATGACA 1668
Db      |||
Qy 1093 rgGlyAlaThrTyrAsnIleIleValGlnAlaLeuLysAspGlnAlaArgHisLysVal 1113
Qy 1669 CCAACGCTATCATTAACCGTTTATATGCGCAAGCAACCGAAGAGAGATCTAGCTAAC 1728
Db      |||
Qy 1113 rgGlnGlnValValThrValGlyAsnSerValAsn--GlnGlyLeuAsnGln----- 1129
Qy 1729 ATTGAGCTGGTGGTGGCCAGCGCAACAGATTGACCATGCTGAGAAAGTTTGAATC 1788
Db      |||
Qy 1130 -----ProThrAspAspSerCysPheAsp 1138
Qy 1789 ATGCTGTGGGACTTCTGATGTGTCGAGAAACGTGGAGAACCCCTTACCAAGGCTGGA 1848
Db      |||
Qy 1138 roTyrThrValSerHisTyrAlaValAlaLysAspGlnThrGlnArgMetSerLysGly 1158
Qy 1849 TGATGTAGATTTGACTTGTGCTGGAGAGAGCAAGCAAGCATGTCATCTCTAGAA 1908
Db      |||
Qy 1158 heLysLeuLeuCysGlnCysLeuGlyPheGlySerGlyHisPheArgCysAspSerSera 1178
Qy 1909 ATGATGCAACGATCAGAGACAAAGCAATCTTATGAATTTGAGACACCTGAGACAA 1968
Db      |||
Qy 1178 rgTyrCysHisAspAsnGly-----ValAsnTyrLysIleGlyGlnLysTyrAspArg 1196
Qy 1969 AGGATATACGAGAAACCTGCTCCAGTGCATCTGCACAGGCAAGCGCCGAGAGAGTGA 2028
Db      |||
Qy 1196 lngLysGlnAsnGlyGlnMetMetSerCysThrCysLeuGlyAsnGlyLysGlnPhe 1216
Qy 2029 AGTGTGAGAGGCAC 2042
Db      |||
Qy 1216 ysCysAspProHis 1220

```

Search completed: January 28, 2006, 02:58:49
 Job time : 44.0689 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 28, 2006, 01:49:35 ; Search time 32.1209 Seconds
(without alignments)
12556.953 Million cell updates/sec

Title: US-09-940-235-12

Perfect score: 3726
Sequence: 1 cgaagaccatcatcgtgtt.....ccttcacgatgtcgttag 2096

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

MODEL=frame+ n2p.model -DEV=xlp
-O=/cgn2.1/USPTO_spool_p/US09940235/rnatc_27012006_144218_27578/app_query.fasta_1.7708
-DB=ptr -QFMT=fastan -SUFRTX=rpr -MINMATCH=0.1 -LOOPEX=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIG=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USERS=US09940235 @CCN_1_1.185 @runatc_27012006_144218_27578 -NCPD=6 -ICPU=3
-NO MMAP -LANG=JUBY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: pir_80:*
2: pir1:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Query length	DB ID	Description
1	1984	53.2	440	1 A22801	streptokinase prec
2	1951.5	52.4	415	1 BZSO	streptokinase A (E)
3	1942	52.1	440	2 S02723	streptokinase G pr
4	1779	47.7	440	2 S02724	streptokinase A pr
5	1673	44.9	414	2 JU0292	streptokinase A pr
6	1670	44.8	440	2 S04168	streptokinase A pr
7	667	17.9	197	2 S53334	streptokinase - St
8	640.5	17.2	2386	1 FNMU	fibronectin precu
9	624.5	16.8	2265	1 FNMU	fibronectin - bovi
10	620	16.6	128	2 S77671	streptokinase A (E)
11	587	15.8	2477	2 S14428	fibronectin precu
12	569	15.3	128	2 S77680	streptokinase A (E)
13	557	14.9	128	2 S77688	streptokinase A (E)
14	554.5	14.9	2481	2 A43908	fibronectin - Afri

15	552	14.8	128	2 S77679	streptokinase A (E)
16	423	11.4	128	2 S77676	streptokinase A (E)
17	419	11.2	128	2 S77673	streptokinase A (E)
18	419	11.2	128	2 S77687	streptokinase A (E)
19	417	11.2	128	2 S77682	streptokinase A (E)
20	416	11.2	128	2 S77678	streptokinase A (E)
21	408	11.0	128	2 S77677	streptokinase A (E)
22	407	10.9	128	2 S77685	streptokinase A (E)
23	401	10.8	128	2 S77683	streptokinase A (E)
24	398	10.7	128	2 S77686	streptokinase A (E)
25	397	10.7	128	2 S77684	streptokinase A (E)
26	396	10.6	128	2 S77681	streptokinase A (E)
27	390	10.5	128	2 S77675	streptokinase A (E)
28	386	10.4	128	2 S77674	streptokinase A (E)
29	226	6.1	1020	2 A29355	streptokinase A (E)
30	182	4.9	190	2 I51279	fibronectin - chic
31	146.5	3.9	2231	2 D71870	hypothetical prote
32	145	3.9	132	1 QOEC8	hypothetical prote
33	138	3.7	1315	2 T28679	fibronogen-binding
34	136.5	3.7	2334	2 S32920	cell wall-associat
35	132	3.5	1385	2 D89824	hypothetical prote
36	131	3.6	1350	2 T10294	hypothetical prote
37	130.5	3.5	940	2 AD1374	invariant surface
38	129.5	3.5	13055	2 T16580	internalin protei
39	128	3.4	2256	2 AD1018	large repetitive p
40	126.5	3.4	1512	2 T14883	hypothetical prote
41	126	3.4	4688	2 P82885	hypothetical prote
42	125.5	3.4	940	2 AB1744	internalin protein
43	125	3.4	1002	2 T02468	hypothetical prote
44	125	3.4	1220	1 DJBEC3	DNA-directed DNA p
45	125	3.4	1417	2 H90670	probable invasiv

ALIGNMENTS

RESULT 1

A22801
Streptokinase precursor - Streptococcus "equisimilis"

C/Species: Streptococcus "equisimilis"

C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004

C/Accession: A22801

R/Malke, H.; Roe, B.; Ferretti, J. J.

Gene 34, 357-362, 1985

A/Title: Nucleotide sequence of the streptokinase gene from Streptococcus equisimilis H

A/Reference number: A22801; MUID:85232082; PMID:2989113

A/Accession type: DNA

A/Residues: 1-440 <MAL>

A/Cross-references: UNIPROT:P00779; UNIPARC:UPI000002BE73; GB:X72832; NID:9407876; PIDN

A/Experimental source: strain H46A

C/Genetics:

A/Gene: skc

C/Superfamily: streptokinase

Alignment Scores:

Pred. No.: 3,29e-126
Score: 1984.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 53.25%
DB: 1
Gaps: 0

US-09-940-235-12 (1-2096) x A22801 (1-440)

QY	588	ATTCCTGACCTGAGTGGCTGCTGACGCTCATCTGTCACACAGCAATGTTGTT	647
DB	27	TTCCTGACCTGAGTGGCTGCTGACGCTCATCTGTCACACAGCAATGTTGTT	46
QY	648	AGCGTTCGCTGCTGCTGTTGAGGGGACCAATCAAGACATTAATTTTGAATC	707
DB	47	TTCCTGACCTGAGTGGCTGCTGACGCTCATCTGTCACACAGCAATGTTGTT	66
QY	708	GATCTAACATCAAGCACTGCTCATGAGGAAAGACAGCAAGGCTTAAGTCAAAATCA	767

```

Db      |||
67  AspleuThrSerArgProAlaHisGlyGlySerThiGluGlnGlyLeuSerProLysSer 86
Qy      768  AAACCAATTTGCTACTGATAGTGGCGCGATGTCACATTAACCTTGAGAAAGCTGACTTA 827
Db      87  LysProPheAlaThrAspSerGlyAlaMetSerHisLeuLeuGlnLysAlaAspLeuLeu 106
Qy      828  AAGGCTATTCAAGAACAAATTGATCGCTAACGTCACAGTAACGACATCTTGAGGTC 887
Db      107  LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 126
Qy      888  ATTGATTTTGAAGCGATGCAACCATTAAGTCAAGCGACGAGCTTACTTGCTGAC 947
Db      127  IleAspPheAlaSerThrAspAlaThrIleThrAspArgAsnGlySerValTyrPheAlaAsp 146
Qy      948  AAAGATGGTTCGGTAACTTGGCCGACCCAACTGTCGAAGAAATTTTGTACCGGACAT 1007
Db      147  LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 166
Qy      1008  GTGCGCGTTAGACCATATAAAGAAAAACCAATACCAAAACCAAGCGAAATCTGTGATGTC 1067
Db      167  ValArgValArgProTyrLysGlyLysProIleGlnAsnGlnAlaLysSerValAspVal 186
Qy      1068  GAAATPACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTGACACGAGCTCAAA 1127
Db      187  GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 206
Qy      1128  GATATGAAGCTATTGAAAACTAGCTATCGGTGACACCATCAATCTGAGAAATTA 1187
Db      207  AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 226
Qy      1188  GCTCAAGCAACAAGCAATTTTAAACAAAAACCAACCGGCTATACGATTTATGAGCTGAC 1247
Db      227  AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluAsp 246
Qy      1248  TCCTCAATCGTCACTGATGACATGACATTTTCCGTACGATTTTCAATGATGACAG 1307
Db      247  SerSerIleValIleThrHisAspAsnAspIlePheAspThrIleLeuProMetAspGlnGlu 266
Qy      1308  TTTACTTACCGTGTAAATAATCGGGAACAAGCTTATAGATCAATTAATAATCTGCTG 1367
Db      267  PheThrTyrArgValLysAsnArgGlnGlnAlaTyrArgIleAsnLysLysSerGlyLeu 286
Qy      1368  AATGAAGAAATTAACAACACTGACTGATCTCTGAGAAATTAATTAATCTGCTTAAAAAGG 1427
Db      287  AsnGlnGlnIleAsnAsnThrAspLeuIleSerGlnLysTyrTyrValLeuLysGly 306
Qy      1428  GAAAGCCGTATGATCCCTTGTATGCGACATCTGAAACGTGTCACCATCAATACGTT 1487
Db      307  GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326
Qy      1488  GATGTCGATACCAACGAATTTGCTAAATAAGTAGACGCTTAAACAGCTACGGAACCTAAC 1547
Db      327  AspValAspThrAsnGlnLeuLeuLysSerGlnGlnLeuThrIleAspGlnAsn 346
Qy      1548  TTAAGCTTCAGAGATTATACATCTCTGTGATAGGCTTAAACTTACATCAACAAATTC 1607
Db      347  LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
Qy      1608  GATGCTTTTGTGATATAGCAATACCTTAACCTGAAAGTAGAGATTAACAGATGAC 1667
Db      367  AspAlaAspGlyLysIleMetAspTyrThrIleThrGlnLysValGlnLysAsnHisAspAsp 386
Qy      1668  ACCAACCGTATACATTAACCTTTATATAGGCAAGCGAACCGGAAGAGATGCTAGCTAC 1727
Db      387  ThrAsnArgIleIleThrValTyrMetGlyLysArgProGlnGlyGlnAsnAlaSerTyr 406
Qy      1728  CATTTAGCT 1736
Db      407  HisLeuAla 409

```

RESULT 2
BZSO

```

streptokinase (EC 3.4.-.-) - Streptococcus sp.
C:Species: Streptococcus sp.
C>Date: 05-Apr-1983 #Sequence_revision 05-Apr-1983 #text_change 09-Jul-2004
C/Accession: A00967
R/Jackson, K.W.; Tang, J.
Biochemistry 21, 6520-6525, 1982
A>Title: Complete amino acid sequence of streptokinase and its homology with serine protease
A/Reference number: A00967; PMID:83127125; PMID:6760891
A/Accession: A00967
A/Molecule type: protein
A/Residues: 1-415 <JAC>
A/Cross-references: UNIPROT:P00779; UNIPARC:UPI0000172BFS
A/Note: 169-Asp and 181-Asp were also found
A/Note: this protein is not a protease, but it activates plasminogen by complexing with C/Superfamily: streptokinase
C/Keywords: hydrolase

Alignment Scores:
Pred. No.: 5,1e-124 Length: 415
Score: 1951.50 Matches: 380
Percent Similarity: 99.22% Conservative: 1
Best Local Similarity: 98.96% Mismatches: 2
Query Match: 52.38% Indels: 1
DB: 1 Gaps: 1

US-09-940-235-12 (1-2096) x BZSO (1-415)
Qy      588  ATTGCTGACCTGATGAGCTGCTAGACCGTCACTCTGTCAACAACGCAATGTTGTT 647
Db      1  IleAlaGlyProGluThrLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
Qy      648  AGCGTGTCTGTACTGTGAGGGGACGATCAAGACATTTAGTCTTAATTTTGAATC 707
Db      21  SerValAlaGlyThrValGlnGlyThrAsnGlnAspIleSerLeuLysPheGluLeu 40
Qy      708  GATCTAAACATCAGACCGTGTCAATGAGGAAAGACAGAGAAAGCTTAAGTCAAAATCA 767
Db      41  AspleuThrSerArgProAlaHisGlyGlyLysThrIleGlnGlyLeuSerProLysSer 60
Qy      768  AAACCAATTTGCTACTGATAGTGGCGCGATGTCACATTAACCTTGAGAAAGCTTACTA 827
Db      61  LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlnLysAlaAspLeuLeu 80
Qy      828  AAGGCTATTCAAGAACAAATTGATCGCTAACGTCACAGTAACGACATCTTGAGGTC 887
Db      81  LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100
Qy      888  ATTGATTTTGAAGCGATGCAACCATTAAGTCAAGCGACGATCTTGCTGAC 947
Db      101  IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
Qy      948  AAAGATGGTTCGGTAACTTGGCCGACCCAACTGTCGAAGAAATTTTGTACGAGCAT 1007
Db      121  LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
Qy      1008  GTGCGCGTTAGACCATATAAAGAAAAACCAATACCAAAACCAAGCGAAATCTGTGATGTC 1067
Db      141  ValArgValArgProTyrLysGlyLysProIleGlnAsnGlnAlaLysSerValAspVal 160
Qy      1068  GATATPACTGTACAGTTTACTCCCTTAAACCCGTGACAGATTTTCAACAGGCTCAAA 1127
Db      161  GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
Qy      1128  GATACTAAGCTATTGAAAACTAGCTATCGGTGACACCATCAATCATCTCAAGATTA 1187
Db      181  LeuThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 200
Qy      1188  GCTCAAGCAACAAGCAATTTTAAACAAAAACCAACCGGCTATACGATTTATGAGCTGAC 1247
Db      201  AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluAsp 220
Qy      1248  TCCTCAATCGTCACTGATGACATGACATTTTCCGTACGATTTTCAATGATGACAG 1307

```

```

Db      221 SerSerIleValIthrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
Qy      1308 TTACTTACCGGTGTAAATAATCGGGAACAAGCTTATAGATCAATAATAAATCGTGTG 1367
Db      241 PheThrTyArgValIlysaAsnArgGlnGlnAlaTyArgIleAsnIlyblySerGlyLeu 260
Qy      1368 AATGAAGAAATAACAACACTGACCTGATCTTGAGAAATAT--TACGTCTTAAAAA 1424
Db      261 AsnGlnGlnIleAsnAsnThrAspLeuIleSerLeuGlnIlyTyArgValIleuIlybly 280
Qy      1425 GGGGAAAAGCCGATATATCCCTTTGATCGGAGTCACTTGAACCTTCAACATCAATAC 1464
Db      281 GlyIlyblyProTyArgProPheAspArgSerHisLeuIlyblyPheThrIleLysGlyr 300
Qy      1485 GTTGATGTGATACCAACGAATGTCTTAAAAAGTAGAGAGCTTTAACAGTAGAGCAAGT 1544
Db      301 ValAspValAspThrAsnGlnLeuLeuIlyblySerGlnIleuLeuThrAlaSerGlnArg 320
Qy      1545 AACTTAGACTTCAGAGATTATACGATCCTCGTATTAAGCTTAACTTACTTACACAAAT 1604
Db      321 AsnLeuAspPheArgAspLeuTyArgProArgAspIlyblyAlaIlyblyLeuTyArgAsn 340
Qy      1605 CTCGATGCTTTGGTATTTATGACACTTATCCTTAACTGAAAAGTAGAGATATCAGAT 1664
Db      341 LeuAspAlaPheGlyIleMetAspTyThrLeuThrGlyIlyblyValGluAspAsnHisAsp 360
Qy      1665 GACACCAACCGTATCAATACCGTTTATATGGGCAAGGACCCGGAAGGAGAAATGTACG 1724
Db      361 AspThrAsnArgIleIleThrValTyArgGlyIlyblyArgProGlnGlyIlyblyAsnLaser 380
Qy      1725 TACCATTTAGCT 1736
Db      381 TyrHisLeuAla 384

RESULT 3
S02723
Streptokinase G precursor - Streptococcus sp. (strain 19908)
C:Species: Streptococcus sp.
C:Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 09-Jul-2004
C:Accession: S02723
R:Walter, F.; Siegel, M.; Malke, H.
Nucleic Acids Res. 17, 1262, 1989
A:Title: Nucleotide sequence of the streptokinase gene from a group-G Streptococcus.
A:Reference number: S02723; MID:89160265; PMID:2922269
A:Accession: S02723
A:Molecule type: DNA
A:Residues: 1-440 <MAL>
A:Cross-references: UNIPROT:P10519; UNIPARC:UPI000013614B; EMBL:X13400; NID:g47095; PDB
C:Genetics:
A:Gene: skg
C:Superfamily: streptokinase
F:1-26/Domain: signal sequence #status predicted <SIG>
F:27-440/Product: streptokinase #status predicted <MAT>

Alignment Scores:
Pred. No.: 2,256-123 Length: 440
Score: 1942.00 Matches: 375
Percent Similarity: 98.43% Conservative: 2
Best Local Similarity: 97.91% Mismatches: 6
Query Match: 52.12% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-12 (1-2096) x S02723 (1-440)
Qy      588 ATTGTCGATCGATGAGCGCTGATGACCGTCAATCTGTCAACAAGCAATGTTGTT 647
Db      27 IleAlaGlyProGlnIlyblyPheAspArgProSerValAsnAsnSerGlnLeuVal 46
Qy      648 ACCGTTGCTGTGATCTGTGAGGGGAGCAATCAAGACATTAGTCTTAATTTTGAATC 707
Db      47 SerValAlaGlyThrValGlnGlyThrAsnGlnAspIleSerLeuIlyblyPheGlnIle 66
Qy      708 GATCTAACATCAAGACCTGCTCTATGAGGAAAGACAGACGAGGCTTAAGTCCAAATCA 767

```

```

Db      67 AspLeuThrSerArgProAlaHisGlyGlySerThrGlnGlnGlyLeuSerProIlyblySer 86
Qy      768 AAACATTGTTGCTACTGATAGTGGCCGATGTCACATTAACCTTGAGAAAGCTGACTTACTA 827
Db      87 LysLeuPheAlaThrAspSerGlyAlaMetProHisLysLeuGlnIlyblyAlaAspLeuLeu 106
Qy      828 AAGCTATTCAAGAACATTTGATGCTAACGCCGACAGTAAACGACGACTTACTTGAAGTC 887
Db      107 LysAlaIleGlnGlnGlnIleuIleAlaAsnValHisSerAsnAspAspTyPheGlnVal 126
Qy      888 ATTGATTTTGGCAAGCATGACCAACATTAATGATCGCAAAACGGCAAGTCTACTTCTGAC 947
Db      127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyIlyblyValTyArgPheAlaAsp 146
Qy      948 AAAGATGTTGGTATCCTTGGCCGACCACTGTCGCAAGAAATTTTGCTAAGCGCAAT 1007
Db      147 LysAspGlySerValThrLeuProIleGlnProValGlnGlnPheLeuLeuIlyblyHis 166
Qy      1008 GTGGCGCTTAGACCATATATTAAGAAACCAATPACAAACCAAGCAAAATCGTTGATGTG 1067
Db      167 ValArgValArgProTyArgGlyIlyblyProValGlnAsnGlnAlaLysSerValAspVal 186
Qy      1068 GAATATACCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTGACACCAAGTCTCAAA 1127
Db      187 GluTyThrValGlnPheThrProLeuAsnProAspAspAspPheArgProAlaLeuIlybly 206
Qy      1128 GATACATAAGCTATTTGAAAACACTAGCTATCGGTGACACCATCACACTTCAAGAAATTACTA 1187
Db      207 AspThrLysLeuLeuIlyblyThrLeuAlaIleGlyAspThrIleThrSerGlnIlyblyLeu 226
Qy      1188 GCTCAGACGACAAACATTTTAAACAAAACCAACGAGCTTATGACATTTAAGAAGTAC 1247
Db      227 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyThrIleTyArgIlyblyAsp 246
Qy      1248 TCCTCAATCGTCACTGATGACAAATGACATTTTCCGTAGATTTTACCAATGATGATCAAG 1307
Db      247 SerSerIleValIthrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
Qy      1308 TTACTTACCGGTGTAAATAATCGGGAACAAGCTTATAGATCAATAATAAATCGTGTG 1367
Db      267 PheThrTyArgValIlysaAsnArgGlnGlnAlaTyArgIleAsnIlyblySerGlyLeu 286
Qy      1368 AATGAAGAAATAACAACACTGACCTGATCTCTGAGAAATTTACGTCCTTAAAAAGGG 1427
Db      287 AsnGlnGlnIleAsnAsnThrAspLeuIleSerGlnIlyblyTyArgValIleuIlyblyGly 306
Qy      1428 GAAAAGCGATGATCCCTTTGATCGGAGTCACTGAAACGTTCACATCAATTAAGTT 1487
Db      307 GluIlyblyProTyArgProPheAspArgSerHisLeuIlyblyPheThrIleLysTyArgVal 326
Qy      1488 GATGTCGATACCAACGAATGCTTAAAAAGTAGAGAGCTTTAACAGTAGAGCAAGTAAAC 1547
Db      327 AspValAsnThrAsnGlnLeuLeuIlyblySerGlnIleuLeuThrAlaSerGlnArgAsn 346
Qy      1548 TTAGACTTCAGAGATTATACGATCTCGTGATTAAGGCTTAATCTTCAACAAATCTC 1607
Db      347 LeuAspPheArgAspLeuTyArgProArgAspIlyblyAlaIlyblyLeuTyArgAsnLeu 366
Qy      1608 GATGCTTTGGTATTAAGCACTATACCTTAACTGAAAAGTAGAGAGATTAACAAGATAC 1667
Db      367 AspAlaPheGlyIleMetAspTyThrLeuThrGlyIlyblyValGluAspAsnHisAspAsp 386
Qy      1668 ACCAAGCGTATCAATACCGTTTATATGGGCAAGGACCCGGAAGGAGAAATGTAGTAC 1727
Db      387 ThrAsnArgIleIleThrValTyArgGlyIlyblyArgProGlnGlyIlyblyAsnAlaSerTy 406
Qy      1728 CATTTAGCT 1736
Db      407 HisLeuAla 409

RESULT 4
S02724

```

streptokinase A precursor - Streptococcus pyogenes (strain SF130/13)
 C/Species: Streptococcus pyogenes
 C/Date: 18-Oct-1989 #sequence_revision 18-Oct-1989 #text_change 09-Jul-2004
 C/Accession: S02724
 R/Walter, F.; Siegel, M.; Malke, H.
 Nucleic Acids Res. 17, 1261, 1989
 A/Title: Nucleotide sequence of the streptokinase gene from a Streptococcus pyogenes type
 A/Reference number: S02724; MUID:89160264; PMID:2646590
 A/Accession: S02724
 A/Molecule type: DNA
 A/Residues: 1-440 <MAL>
 A/Cross-references: UNIPROT:P10520; UNIPARC:UPI000017021D; EMBL:X13399; MID:947435; PIDD
 C/Genetics:
 A/Gene: ska
 C/Superfamily: streptokinase
 F:1-26/Domain: signal sequence #status predicted <Sig-
 F:27-440/Product: streptokinase #status predicted <MAT>

Alignment Scores:
 Pred. No.: 2,266-112 Length: 440
 Score: 1779.00 Matches: 338
 Percent Similarity: 93.73% Conservative: 21
 Best Local Similarity: 88.25% Mismatches: 24
 Query Match: 47.75% Indels: 0
 DB: 2 Gaps: 0

US-09-940-235-12 (1-2096) x S02724 (1-440)

QY 588 ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCAATTTGGTTG 647
 |||||
 Db 27 IleaIaGIyTyrGIyTyrleuProAspArgProProlleAsnSerGlnleuValVal 46
 |||||
 QY 648 AGCGTTGCTGCTGCTGAGGGGAGCGATCAAGACATTAGCTTAAATTTTGAATC 707
 |||||
 Db 47 SerMetlaGIyIleValGlnGlyThrAspLysValPheIleAsnhePheGluIle 66
 |||||
 QY 708 GATCAATCATGACACCTGCTCATGAGGAGAAAGACAGGCTTAAGTCCAAATCA 767
 |||||
 Db 67 AspLeuThrSerGlnProIleAsnGlyLysThrGlnGlnGlyLeuSerProLysSer 86
 |||||
 QY 768 AAACCATTTGCTAGTATGATGCGCGAGATGTCAATTAACCTTGAGAACCTGACTTA 827
 |||||
 Db 87 LysProPheAlaThrAspAsnGlyLysMetProHisLysLeuGlnLysAlaAspLeu 106
 |||||
 QY 828 AAGGATTTCAAGAACATTTGCTTACGTCACAGTACAGACGACTTGTGAGTC 887
 |||||
 Db 107 LysAlaIleGlnLysGlnleuIleAlaAsnValHisSerAspAspGlyTyrPheGluVal 126
 |||||
 QY 888 ATTGATTTGCAAGGAGATGACCATTTACTGATTCGAAGGAGGCTTACTTGTGAC 947
 |||||
 Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146
 |||||
 QY 948 AAAGATGCTTGGTACCTTGGCCGACCCAACTGTCAGAAATTTTGTCAAGCGGACAT 1007
 |||||
 Db 147 LysAspGlySerValThrleuProThrGlnProValGlnGlnPheLeuValLysGlyHis 166
 |||||
 QY 1008 GTGCGCGTTAGACCATTAAGAAAAACCAATTAACCAACGGAATCTGTTAGTGG 1067
 |||||
 Db 167 ValArgValArgProTyrLysGlnLysProValGlnAsnGlnAlaLysSerValAspVal 186
 |||||
 QY 1068 GAATATCTGTCAGATTACTCCCTTAACCCCTGATGAGATTGTCAGACGCTCCAA 1127
 |||||
 Db 187 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 206
 |||||
 QY 1128 GATACTAAGCTATTGAAAACTAGCTATCGGTGACACATCATCTCAAGAAATTA 1187
 |||||
 Db 207 AspThrLysLeuLeuLysThrleuAlaIleGlyAspThrIleThrSerGlnGlnleuLeu 226
 |||||
 QY 1188 GCTCAAGCAAAAGATTTTAAACAAAACCAACGAGGCTATACGATTTATGAACGTGAC 1247
 |||||
 Db 227 AlaGlnIleAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGlnArgAsp 246
 |||||
 QY 1248 TCCCAATCGTCACATGACAAATGACATTTTCCGATGACATTTTACCATGATGACAGAG 1307
 |||||

Db 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
 |||||
 QY 1308 TTACTTACCGGTGTTAAAAATCGGGAACAGCTTATAGATCAATTAATAATCGTCTG 1367
 |||||
 Db 267 PheThrTyrHisValLysAsnArgGlnGlnAlaTyrGlnIleAsnProLysThrGlyIle 286
 |||||
 QY 1368 AATGAGAAATTAACAACAGCTGACCTGATCGAGAAATTTATACGTCCTTAATAAAGG 1427
 |||||
 Db 287 LysGlnLysThrAsnSerThrAspLeuValSerGlnLysTyrTyrValLeuLysGlnGly 306
 |||||
 QY 1428 GAAAAGCGGTATGATCCCTTGTATCGAGTCACTTGAACCTGTCACATCAATTAAGTT 1487
 |||||
 Db 307 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326
 |||||
 QY 1488 GATGTCGATTCGAACGAAATGCTTAAAGTGAAGAGCTTTTACAGCTAGGAAAGTAC 1547
 |||||
 Db 327 AspValAsnThrAsnGlnleuLeuLysSerGlnGlnleuLeuThrAlaSerGlnGlyAsn 346
 |||||
 QY 1548 TTAGACTTCAGAGATTTATACGATCCCGTATGAGCTTAACTCTCAACAATCTC 1607
 |||||
 Db 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLysTyrAsnAsnleu 366
 |||||
 QY 1608 GATGCTTTTGGTATTTATGACACTTACTTACTGAAAGAGTAAAGAGATTAATCAAGATGAC 1667
 |||||
 Db 367 AspAlaPheAspIleMetAsnTyrThrLeuThrGlyLysValGlnAspAsnHisAspLys 386
 |||||
 QY 1668 ACCAACCGTATCATTAACCGTTTATATGAGGAGACCCGAGAGAGAAATGCTAGTAC 1727
 |||||
 Db 387 AsnAsnArgValAlaThrValTyrMetGlyLysArgProLysGlyAlaLysGlySerTyr 406
 |||||
 QY 1728 GATTAAGCT 1736
 |||||
 Db 407 HisLeuAla 409
 |||||

RESULT 5
 J0292
 streptokinase - Streptococcus pyogenes
 C/Species: Streptococcus pyogenes
 C/Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
 C/Accession: A43867; J0292
 R/Ohkuni, H.; Todome, Y.; Suzuki, H.; Mizuse, M.; Kotani, N.; Horuchi, K.; Shikama, N.;
 Infect. Immun. 60, 278-283, 1992
 A/Title: Immunological studies and complete amino acid sequence of the streptokinase fr
 A/Reference number: A43867; MUID:92104686; PMID:1370275
 A/Accession: A43867
 A/Molecule type: protein
 A/Residues: 1-414 <OHK>
 A/Cross-references: UNIPROT:Q57391; UNIPARC:UPI0000175C82
 A/Experimental source: M type 12 strain A374
 A/Note: sequence extracted from NCBI backbone (NCBIP:74592)
 C/Superfamily: streptokinase

Alignment Scores:
 Pred. No.: 3,196-105 Length: 414
 Score: 1673.00 Matches: 324
 Percent Similarity: 90.60% Conservative: 23
 Best Local Similarity: 84.60% Mismatches: 36
 Query Match: 44.90% Indels: 0
 DB: 2 Gaps: 0

US-09-940-235-12 (1-2096) x J0292 (1-414)

QY 588 ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCAATTTGGTTG 647
 |||||
 Db 1 IleaIaGIyProGluTyrleuLeuAspArgProSerValAsnSerGlnleuValVal 20
 |||||
 QY 648 AGCGTTGCTGCTGCTGAGGGGAGCAATCAAGACATTAGCTTAAATTTTGAATC 707
 |||||
 Db 21 SerValAlaGIyThrValGlnGlyThrAsnGlnGlyIleSerLeuLysPhePheGluIle 40
 |||||
 QY 708 GATCAATCATGACACCTGCTCATGAGGAGAAAGACAGGCTTAAGTCCAAATCA 767
 |||||

```
Dh 41 AspleuThSerArpProAlaHISGLyGlyuThrGluGlnGlyLeuSerProLySer 60
Qy 768 AAACCATTTGCTACTGATAGTGGCGCATGTCACATTAACCTTGAGAAAGTGACTTACTA 827
Db 61 LysProPheAlaThrAsnSerSerAlaMetProHISlySLeuGluValAlaAspLeu 80
Qy 828 AAGCTATTTCAGAAACAATTGATGCTTAACCTGCACAGTAAAGACGACTTTGAGGTC 887
Db 81 LysAlaAlaGlnGlnGlnLeuAlaAlaAsnValHISerAsnAsnAspGlyTYrPheGluVal 100
Qy 888 ATTGATTTTGGACCGCATGCAACCATTAATCTGATGAAACGGCAGGCTTACTTTGCTGAC 947
Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyuValTYrPheAlaAsp 120
Qy 948 AAAGATGGTTCGGTAACTTGGCCGACCCGACCTGTCCAGAAATTTTGGTAAAGGACAT 1007
Db 121 ArgAspAspSerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHIS 140
Qy 1008 GTGGCGGTTAGACCATATTAAGAAAAACAATACAAAACAAGCGAAATCTGTGATGTG 1067
Db 141 ValArgValArgProTYrGlnProLYuAlaValHISAsnSerAlaGluArgValAlaAsnVal 160
Qy 1068 GAATATACCTGTACATTACTCCCTTAAACCTGATGACGATTTTCAGACCGTCTCAA 1127
Db 161 AsnTYrGluValSerPheValSerGluThrGlyAsnLeuAspPheThrProSerLeuLYs 180
Qy 1128 GATACATACGCTATTAAGAAACCTAGCTATGGTGAACCATCAATCAATCTCAAGATTACTA 1187
Db 181 GluArgTYrHISLeuThrThrLeuAlaValGlyAspSerLeuSerSerGlnGluLeuAla 200
Qy 1188 GCTCAAGACCAAAAGATTTTAAACAAAACACACCGCTATACGATTTTATGACGTGAC 1247
Db 201 AlaIleAlaGlnPheIleLeuSerLYsGlnHISProAspTYrIleIleThrLYsArgAsp 220
Qy 1248 TCTCAATGCTCATCTATGACATGACATTTTCCGATCGAATTTTACCAATGATCAAGAG 1307
Db 221 SerSerIleValThrHISAspAsnAspIlePheArgThrIleLeuProMetArgGlnGlu 240
Qy 1308 TTTACTTACCGTGTAAAAATCGGGAACAAGCTTATAGATCATTAATAAAATCTGGTCTG 1367
Db 241 PheThrTYrHISIleLYsAspArgGlnGlnAlaTYrLYsAlaAsnSerLYsThrGlyIle 260
Qy 1368 AATGAGAATAATAACAACGACGACCTGATCTGAGAAATATTGATGCTTAAAAAAAGGG 1427
Db 261 ValGluLYrThrAsnAsnThrAspLeuIleSerGlnLYsTYrTYrValIleLYsLYsGly 280
Qy 1428 GAAACCCGATATGATCCCTTGAATCGCATGTCATTTGAACGTGTCACATCAATAAGTT 1487
Db 281 GluGluProTYrAspProPheAspArgSerHISLeuLYsLeuPheThrIleLYsTYrVal 300
Qy 1488 GATGTGATACCAACGAAATTTGCTAAAGAGGACGCTTTAACAGCTTACGGAACGTGAC 1547
Db 301 AspValAspThrAsnGlnLeuLeuLYsSerGlnGlnLeuLeuThrAlaSerGlnLYsArgAsn 320
Qy 1548 TTAGATTTGAGAGATTATACATCTCCGATGATAGGCTAAACCTCAACAACATCTC 1607
Db 321 LeuAspPheArgAspLeuTYrAspProArgAspLYsAlaLYsLeuLeuTYrAsnAsnLeu 340
Qy 1608 GATGCTTTGGTATTATGACTATACCTTAACCTGAAAGAGTAATACAGATGAC 1667
Db 341 AspAlaPheGlyIleMetAspTYrThrLeuThrGlyLYsValGluAspAsnHISAsnAsp 360
Qy 1668 ACCCAACCGTATCAAAACGTTTATATGAGGACAGACCCGAAAGAGAAATGCTAGCTAC 1727
Db 361 ThrAsnArgIleIleThrValTYrMetGlyLYsArgProGluGlnGluAsnAlaSerTYr 380
Qy 1728 CATTTAGCT 1736
Db 381 HISLeuAla 383
```

RESULT 6
S04168
streptokinase A precursor - Streptococcus pyogenes (strain NZ131)

```
C:Species: Streptococcus pyogenes  
C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #ext_change 09-Jul-2004  
C:Accession: S04168  
R:Huang, T.T.; Malke, H.; Ferretti, J.J.  
Mol. Microbiol. 3, 197-205, 1989  
A:Title: The streptokinase gene of group A streptococci: cloning, expression in Escheri-  
A:Reference number: S04168; MUID:89343623; PMID:2668686  
A:Accession: S04168  
A:Molecule type: DNA  
A:Residues: 1-440 <HUA>  
A:Cross-references: UNIPROT:O57391; UNIPARC:UPI0000BD04A; EMBL:X51517; NID:G47437; PTD  
F:1-26/Domain: signal sequence #status predicted <SIG>  
F:27-440/Product: streptokinase A #status predicted <MAT>  
  
Alignment Scores:  
Pred. No.: 5,13e-105 Length: 440  
Score: 1670.00 Matches: 324  
Percent Similarity: 90.34% Conservative: 22  
Best Local Similarity: 84.60% Mismatches: 37  
Query Match: 44.82% Indels: 0  
Gaps: 0  
  
US-09-940-235-12 (1-2096) x S04168 (1-440)  
  
Qy 588 ATTGTGACCTGATAGTGGCTGTGATGACCGTTCATCTGTCACAAACGCAATGTGTTGTT 647  
Db 27 IleaGlyTYrGlnTYrLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 46  
Qy 648 AGCGTGTGCTACTGTTGAGGGGACGAATCAAGCATTTAGCTTAAATTTTGAATC 707  
Db 47 SerValaGlyThrValGlnGlyThrAsnGlnGluIleSerLeuLYsPhePheGluIle 66  
Qy 708 GATCAACATCACGACGCTCATGAGGAAAGACAGAGCAAGGCTTAAGTCAAAATCA 767  
Db 67 AspleuThrSerArpProAlaHISGLyGlyuThrGluGlnGlyLeuSerProLYsSer 86  
Qy 768 AAACCATTTGCTACTGATAGTGGCGCATGTCACATTAACCTTGAGAAAGTGACTTACTA 827  
Db 87 LysProPheAlaThrAspLYsGlyAlaMetSerHISlySLeuGluValAlaAspLeu 106  
Qy 828 AAGCTATTTCAGAAACAATTGATGCTTAACCTGCACAGTAAAGACGACTTGTGAGGTC 887  
Db 107 LysAlaIleGlnGlnGlnLeuAlaAlaAsnValHISerAsnAspGlyTYrPheGluVal 126  
Qy 888 ATTGATTTTGGACCGCATGCAACCATTAATCTGATGAAACGGCAGGCTTACTTTGCTGAC 947  
Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyuValTYrPheAlaAsp 146  
Qy 948 AAAGATGGTTCGGTAACTTGGCCGACCCGACCTGTCCAGAAATTTTGGTAAAGGACAT 1007  
Db 147 LysAspAspSerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHIS 166  
Qy 1008 GTGGCGGTTAGACCATATTAAGAAAAACAATACAAAACAAGCGAAATCTGTGATGTG 1067  
Db 167 ValArgValLYsProTYrGlnProLYuAlaValHISAsnSerAlaGluArgValAlaAsnVal 186  
Qy 1068 GAATATACCTGTACATTACTCCCTTAAACCTGATGACGATTTTACGACCGGCTCAA 1127  
Db 187 AsnTYrGluValSerPheValSerGluThrGlyAspLeuAspPheThrProLeuLeuArg 206  
Qy 1128 GATACATACGCTATTAAGAAACCTAGCTATGGTGAACCATCAATCAATCTCAAGATTACTA 1187  
Db 207 AsnGlnTYrHISLeuThrThrLeuAlaValGlyAspSerLeuSerSerGlnGluLeuAla 226  
Qy 1188 GCTCAAGACCAAAAGATTTTAAACAAAACACACCGGCTATACGATTTATGACGTGAC 1247  
Db 227 AlaIleAlaGlnPheIleLeuSerLYsHISProAspTYrIleIleThrLYsArgAsp 246  
Qy 1248 TCTCAATGCTCATGATGACATGACATTTTCCGATGATTTTACCAATGATCAAGAG 1307
```

```

Db      247 SerSerIleValThrHisAspAsnAepIlePheArgThrIleLeuProMetAspGlnGlu 266
Oy      1308 TTACTTACCGTGTAAATAATCGGGAACAAGCTTATAGATCAATAAATAATCGGTCTG 1367
Db      267 PheThrYHisIleIleLysAspArgGlnGlnAlaTyrlLysAlaAsnSerIleStrnGlyIle 286
Oy      1368 AATGAGAATAAACAACAACATGACCTGATCTCTGAGAAATATTATAGTCTTAAAGG 1427
Db      287 GlnGlnLysThrAsnAsnThrAspLeuIleSerGlnLysTyrlTyrlValLeuLysGly 306
Oy      1428 GAAAAGCCGTATGATCCCTTGTATGCGAGTCACTTGAACCTGTCACCAATACGTT 1487
Db      307 GlnLysProIleYAspProPheAspSerHisLeuLysLeuPheThrIleAsnTyrlVal 326
Oy      1488 GATGTCGATACCAACGAATTTGCTTAAAGAGACGCTCTTAACAGTACGGAACGTAC 1547
Db      327 AspValAsnThrAsnLysLeuLeuLysSerGlnGlnLeuLeuThrAlaSerGlnLysAsn 346
Oy      1548 TTGACCTTCAGAGATTTATACGATCTCTCGTATTAAGCTTAACTACTTACAACATCTC 1607
Db      347 LeuAspPheArgAspLeuTyrlAspProArgAspLysAlaLysLeuLeuTyrlAsnLeu 366
Oy      1608 GATGCTTTGGTATTTATGACATATACCTTAACTGGAAGATGAGATTAATCAAGATGAC 1667
Db      367 AspAlaPheGlyIleMetAspTyrlThrLeuThrGlnLysValGlnAspAsnHisAspAsp 386
Oy      1668 ACCAACCGTATCATACCGTTTATATAGGCAAGCGAACCGGAAGAGAAATGCTAGTAC 1727
Db      387 ThrAsnAlaGlyIleIleThrValIleTyrlMetGlyLysArgProGlnGlnLysAlaSerTyrl 406
Oy      1728 CATTTAGCT 1736
Db      407 HisLeuAla 409

```

RESULT 7

```

S53334
strepStokinase - Streptococcus sp.
C/Species: Streptococcus sp.
C/Date: 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change 09-Jul-2004
C/Accession: S53334
R/Shi, G.Y.; Chang, B.I.; Chen, S.M.; Wu, D.H.; Wu, H.L.
Biochem. J. 304, 235-241, 1994
A/Title: Function of streptokinase fragments in plasminogen activation.
A/Reference number: S53334; MUID:95091634; PMID:798939
A/Accession: S53334
A/Status: preliminary
A/Molecule type: protein
A/Residues: 1-197 <SHI>
A/Cross-references: UNIPROT:Q7ML15; UNIPARC:UPI000017AC2C

```

Alignment Scores:

```

Pred. No.: 2.25e-37 Length: 197
Score: 667.00 Matches: 167
Percent Similarity: 44.71% Conservative: 2
Best Local Similarity: 44.18% Mismatches: 1
Query Match: 17.90% Indels: 208
DB: 2 Gaps: 9

```

US-09-940-235-12 (1-2096) x S53334 (1-197)

```

Oy      588 ATTGCTGACCTGATGCGTGTGCTAGACCGTCCATCTGTCAACAACAGCAATGTTGTT 647
Db      1 IleaIaGlyProGlnIlePheLeuAspArgProSerValAsn-----14
Oy      648 AGCGTGTGCTGATCTTTGAGGGGAGCAATCAAGATTAAGTCTTAATTTTGAATC 707
Db      14 -----14
Oy      708 GATCTTAACATCAGACCTGCTCTATGAGAGAAAGACAGACAGGCTTAAGTCCAAATCA 767
Db      15 -----Ser 15
Oy      768 AAACATTTGTACTGATAGTGGCGGATGTCACATAACTTGAGAAAGCTGACTTACTA 827

```

```

Db      16 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlnLysAlaAspLeuLeu 35
Oy      828 AAGCTTTTCAAGAAACAATGATGCTTAAGCTGCAAGTACAGACACTTGTGAGGTC 887
Db      36 LysAlaIleGlnIlePheGlnIleLeu-----43
Oy      888 ATGATTTTTCAGAGCATGCAACCATTAATGATGCAAGGCAAGGCAAGTCTTGTGATGTC 947
Db      44 -----AsnGlnLysValTyrlPheAlaAsp 51
Oy      948 AAGATGTTTCGTTAATCTTCCGACCCCAACCTGTCCAAGAAATTTGCTAAGCGACAT 1007
Db      52 LysAspIleSerValThr-----57
Oy      1008 GTGCGCGTTAGACCATTAATAAGAAAACCAATACAAACCAAGGAATCTGTGATGTC 1067
Db      58 -----GlnLysProIleGlnAsnGlnAlaLysSerValAspVal 70
Oy      1068 GAATATACCTGTACAGTTACTCCCTTAAACCTGATGACGATTCAGACCAAGTCTGAAA 1127
Db      71 Glu-----71
Oy      1128 GATACATAGCTATTTGAAACACTAGCTATGCTGACACATCATCTCAAGATTACTA 1187
Db      71 -----71
Oy      1188 GCTTCACGACAAAGCATTTTAAACAAACCAACCCAGCGTATACGATTTATGAACGTGAC 1247
Db      72 -----AsnHisProGlnTyrlThrIleTyrlGlnArgAsp 82
Oy      1248 TCCGTACATGTCATCATGACATGACATTTTCGTCAGATTTTAAACGATGCAAGAG 1307
Db      83 SerSerIleVal-----ThrLeuProMetAspGlnGlu 94
Oy      1308 TTACTTACCGTGTAAATAATCGGGAACAAGCTTATAGATCAATAAATAATCGGTCTG 1367
Db      95 PheThrTyrl-----SerGlyLeu 100
Oy      1368 AATGAGAATAAACAACAACACTGACCTGATCTCTGAGAAATATTATGCTTAAAGG 1427
Db      101 AsnGlnGlnIleAsnAsnThrAspLeuIleSer-----111
Oy      1428 GAAAAGCCGTATGATCCCTTGTATGCGAGTCACTTGAACCTGTCACCATCAATAGTT 1487
Db      112 -----TyrlVal 113
Oy      1488 GATGTCGATACCAACGAATTTGCTTAAAGAGAGAGCTTTAACAAGCTAGCAAGTAC 1547
Db      114 AspValAspThrAsnGlnLysLeuLysSerGlnGlnLeuLeuThrAlaSerGlnLysAsn 133
Oy      1548 TTGACCTTCAGAGATTTATACGATCTCTGTGATTAAGGCTTAACTCTTACAACATCTC 1607
Db      134 -----AspLeuTyrlAspProArgAspLysAlaLysLeuLeuTyrlAsnAsnLeu 149
Oy      1608 GATGCTTTGGTATTTATGACATATACCTTAACTGGAAGATGAGATTAATCAAGATGAC 1667
Db      150 AspAlaPheGlyIleMet-----155
Oy      1668 ACCAACCGTATCATACCGTTTATATAGGCAAGGACCCGGAAGAGAAATGCT 1721
Db      156 -----IleIleThrValIleTyrlMetGlyLysArgProGlnGlnLysAlaAsnAla 170

```

RESULT 8

FNUH

fibronectin precursor [validated] - human
N:Alternate names: fibronectin splice form ED-A

C/Species: Homo sapiens (man)
C/Date: 27-Nov-1985 #sequence_revision 31-Mar-1993 #text_change 31-Dec-2004
C/Accession: A26460; A26284; S03917; A24854; A24476; A91008; A93529; A21011; A90495; A22
R/Dean, D.C.; Bowlus, C.L.; Bourgeois, S.
Proc. Natl. Acad. Sci. U.S.A. 84, 1876-1880, 1987
A/Title: Cloning and analysis of the promoter region of the human fibronectin gene.

A:Reference number: A26460; MUID:87175578; PMID:3031656
 A:Accession: A26460
 A:Molecule type: DNA
 A:Residues: 1-49 <DEA>
 A:Cross-references: UNIPROT:P02751; UNIPROT:Q14327; UNIPARC:UPI000016A926; GB:M15801; NT
 R:Oldberg, A.; Ruoslahti, E.
 J. Biol. Chem. 261, 2113-2116, 1986
 A:Title: Evolution of the fibronectin gene.
 A:Reference number: A26284; MUID:86111901; PMID:3003095
 A:Accession: A26284
 A:Molecule type: DNA
 A:Residues: 1447-1540 <OLD>
 A:Cross-references: UNIPARC:UPI000011237; GB:M12549; NID:g182668
 A>Note: The authors translated the codon TTC for residue 1494 as Gju
 R:Paolletta, G.; Henchcliffe, C.; Sebastio, G.; Baralle, F.E.
 Nucleic Acids Res. 16, 3545-3557, 1988
 A:Title: Sequence analysis and in vivo expression show that alternative splicing of ED-B
 A:Reference number: S00846; MUID:88233940; PMID:3375063
 A:Accession: S03917
 A:Molecule type: DNA
 A:Residues: 1594-1767, 'V', 1769-1783 <PAO>
 A:Cross-references: UNIPARC:UPI000017432B; EMBL:X07718; NID:g31402
 A>Note: The authors translated the codon AAC for residue 1631 as Asp
 R:Vibe-Pedersen, K.; Magnusson, S.; Baralle, F.E.
 FEBS Lett. 207, 287-291, 1986
 A:Title: Donor and acceptor splice signals within an exon of the human fibronectin gene;
 A:Reference number: A24854; MUID:87030929; PMID:3770201
 A:Accession: A24854
 A:Molecule type: DNA
 A:Residues: 1992-2147 <VIB>
 A:Cross-references: UNIPARC:UPI000017432C; GB:X04530; NID:g31436
 R:Gutman, A.; Yamada, K.M.; Kornblith, A.
 FEBS Lett. 207, 145-148, 1986
 A:Title: Human fibronectin is synthesized as a pre-propolypeptide.
 A:Reference number: A24476; MUID:87030890; PMID:3770189
 A:Accession: A24476
 A:Molecule type: not compared with conceptual translation
 A:Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-14, 'O', 16-38 <GUT>
 A:Cross-references: UNIPARC:UPI000017432D
 R:Kornblith, A.R.; Umezawa, K.; Vibe-Pedersen, K.; Baralle, F.E.
 EMBO J. 4, 1755-1759, 1985
 A:Title: Primary structure of human fibronectin: differential splicing may generate at l
 A:Reference number: A91008; MUID:85284965; PMID:2992339
 A:Accession: A91008
 A:Molecule type: nucleic acid sequence not shown
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 32-1344, 1346-2080, 2112-2386 <KOR>
 A:Cross-references: UNIPARC:UPI000017432E; UNIPARC:UPI000017432F; GB:X02761
 R:Kornblith, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
 Nucleic Acids Res. 12, 5853-5868, 1984
 A:Title: Human fibronectin: cell specific alternative mRNA splicing generates polypeptid
 A:Reference number: A93529; MUID:84272258; PMID:6462919
 A:Accession: A93529
 A:Molecule type: mRNA
 A:Residues: 973-2080, 2112-2386 <KOR>
 A:Cross-references: UNIPARC:UPI000017432F; UNIPARC:UPI0000174330; GB:X00739
 R:Oldberg, A.; Linney, E.; Ruoslahti, E.
 J. Biol. Chem. 258, 10193-10196, 1983
 A:Title: Molecular cloning and nucleotide sequence of a cDNA clone coding for the cell a
 A:Reference number: A21011; MUID:83290929; PMID:6688418
 A:Accession: A21011
 A:Molecule type: mRNA
 A:Residues: 1434-1537 <OL2>
 A:Cross-references: UNIPARC:UPI0000174331; GB:X00055; NID:g182680; PIDN:AAA52459.1; PID:
 R:Bernard, M.P.; Kolbe, M.; Weil, D.; Chu, M.L.
 Biochemistry 24, 2698-2704, 1985
 A:Title: Human cellular fibronectin: comparison of the carboxyl-terminal portion with re
 A:Reference number: A90495; MUID:85280409; PMID:2992573
 A:Accession: A90495
 A:Molecule type: mRNA
 A:Residues: 1594-2386 <BER>
 A:Cross-references: UNIPARC:UPI000016A928; GB:M10905; NID:g182696; PIDN:AAA52462.1; PID:

R:Umezawa, K.; Kornblith, A.R.; Baralle, F.E.
 FEBS Lett. 186, 31-34, 1985
 A:Title: Isolation and characterization of cDNA clones for human liver fibronectin.
 A:Reference number: A22245; MUID:85231203; PMID:2985004
 A:Accession: A22245
 A:Molecule type: mRNA
 A:Residues: 1948-2067 <UME>
 A:Cross-references: UNIPARC:UPI00000046A; GB:M27589; NID:g182705; PIDN:AAA52465.1; PID
 A:Accession: B22245
 A:Molecule type: mRNA
 A:Residues: 1975-1991, 2017-2039 <UM2>
 A:Cross-references: UNIPARC:UPI0000174332; UNIPARC:UPI0000174333; GB:M27590
 R:Seikiuchi, K.; Kloe, A.M.; Kurachi, K.; Yoshitake, S.; Hakomori, S.
 Biochemistry 25, 4936-4941, 1986
 A:Title: Human liver fibronectin complementary DNAs: identification of two different me
 A:Reference number: 152394; MUID:87026578; PMID:3021206
 A:Accession: 165273
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1978-1990, 2016-2018, 'N', 2020-2081, 2113-2127 <SEK>
 A:Cross-references: UNIPARC:UPI000006804C; GB:M14060; NID:g182701; PIDN:AAA52464.1; PID
 R:Kornblith, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
 Proc. Natl. Acad. Sci. U.S.A. 80, 3218-3222, 1983
 A:Title: Isolation and characterization of cDNA clones for human and bovine fibronectin
 A:Reference number: A21165; MUID:83221567; PMID:6304699
 A:Accession: A21165
 A:Molecule type: mRNA
 A:Residues: 2291-2386 <KOR>
 A:Cross-references: UNIPARC:UPI0000174334; GB:X00799; NID:g182681; PIDN:AAA52460.1; PID
 R:Garcia-Pardo, A.; Pearlsstein, E.; Frangione, B.
 J. Biol. Chem. 258, 12670-12674, 1983
 A:Title: Primary structure of human plasma fibronectin
 A:Reference number: A92398; MUID:84032463; PMID:6630202
 A:Accession: A92398
 A:Molecule type: protein
 A:Residues: 32-47, 'C', 49-51, 'S', 53-72, 'A', 74-290 <GAR1>
 A:Cross-references: UNIPARC:UPI0000174335
 R:Garcia-Pardo, A.; Gold, L.I.
 Arch. Biochem. Biophys. 304, 181-188, 1993
 A:Title: Further characterization of the binding of fibronectin to gelatin reveals the
 A:Reference number: S34791; MUID:93312001; PMID:8333285
 A:Accession: S34791
 A:Molecule type: protein
 A:Residues: 291-300, 551-560 <GAR2>
 A:Cross-references: UNIPARC:UPI0000174336; UNIPARC:UPI0000174337
 R:Griffin, C.A.; Calaycay, J.; Shively, J.E.; Smith, R.L.
 Thromb. Res. 43, 469-477, 1986
 A:Title: Two plasma fibronectin fragments with different gelatin-binding properties.
 A:Reference number: A60904; MUID:87019725; PMID:3532418
 A:Accession: A60904
 A:Molecule type: protein
 A:Residues: 293-301 <GR1>
 A:Cross-references: UNIPARC:UPI0000174338
 R:Calaycay, J.; Pande, H.; Lee, T.; Borel, L.; Stirl, A.; Shively, J.E.; Zardi, L.
 J. Biol. Chem. 260, 12136-12141, 1985
 A:Title: Primary structure of a DNA- and heparin-binding domain (domain III) in human p
 A:Reference number: A23901; MUID:86008277; PMID:3900070
 A:Accession: A23901
 A:Molecule type: protein
 A:Residues: 616-677, 'Q', 679-703, 'PT', <CAN>
 A:Cross-references: UNIPARC:UPI0000174339
 R:Pierschbacher, M.D.; Ruoslahti, E.; Sundelin, J.; Lind, P.; Peterson, P.A.
 J. Biol. Chem. 257, 9593-9597, 1982
 A:Title: The cell attachment domain of fibronectin. Determination of the primary structu
 A:Reference number: A92386; MUID:82265604; PMID:7050098
 A:Accession: A92386
 A:Molecule type: protein
 A:Residues: 1441-1548 <PIE>
 A:Cross-references: UNIPARC:UPI0000141CD5
 A>Note: residues 1524-1527 are responsible for the cell-binding activity
 R:Garcia-Pardo, A.; Roaagno, A.; Frangione, B.
 Biochem. J. 241, 923-928, 1987
 A:Title: Primary structure of human plasma fibronectin. Characterization of a 38 kDa dom

A:Reference number: A32517; MUID:87241275; PMID:3593230

A:Accession: A32517

A:Molecule type: protein

A:Residues: 1589-1630, 'T', 1722-2058 <GAR3>

A:Cross-references: UNIPARC:UPI000017433A

R.Tresselt, T.; McCarty, J.B.; Calaycay, J.; Lee, T.D.; Legesse, K.; Shively, J.E.; Pand

Biochem. J. 274, 731-738, 1991

A:Title: Human plasma fibronectin. Demonstration of structural differences between the A

A:Reference number: S14357; MUID:91190085; PMID:2012601

A:Accession: S14357

A:Molecule type: protein

A:Residues: 1614-1630, 'T', 1722-2081, 2113-2244 <RBE>

A:Cross-references: UNIPARC:UPI000017433B

R.Garcia-Pardo, A.; Pearlstein, E.; Frangione, B.

J. Biol. Chem. 260, 10320-10325, 1985

A:Title: Primary structure of human plasma fibronectin. Characterization of a 31,000-dal

A:Reference number: A23891; MUID:85261459; PMID:4019516

A:Accession: A23891

A:Molecule type: protein

A:Residues: 2071-2080, 2112-2356 <GAR4>

A:Cross-references: UNIPARC:UPI000017433C; UNIPARC:UPI000017433D

C:Comment: The extra domain and connecting strand 3 are subject to developmental and tis

C:Comment: The cellular and plasma fibronectins are high molecular weight glycoproteins,

action, and transformation.

C:Genetics:

A:Gene: GDB:FNI

A:Cross-references: GDB:119135; OMIM:135600

A:Map position: 2q34-2q34

A:Intron: 49/3; 1266/1; 1357/1; 1447/1; 1487/1; 1541/1; 1631/1; 1721/1; 1991/1; 2145/1

C:Superfamily: fibronectin type I repeat homology; fibronectin type II repeat homology;

C:Keywords: acute phase; alternative splicing; cell adhesion; collagen binding; duplicat

F:1-26/Domain: signal sequence #status predicted <SIG>

F:27-31/Domain: propeptide #status predicted <PPO>

F:32-2386/Product: fibronectin #status experimental <MAM>

F:52-87/Domain: fibrin and heparin binding <FHB>

F:97-135/Domain: fibronectin type I repeat homology <1F1>

F:141-119/Domain: fibronectin type I repeat homology <1F2>

F:186-225/Domain: fibronectin type I repeat homology <1F3>

F:231-270/Domain: fibronectin type I repeat homology <1F4>

F:308-608/Domain: collagen binding <CBR>

F:308-342/Domain: fibronectin type I repeat homology <1F6>

F:360-401/Domain: fibronectin type II repeat homology <2F1>

F:420-461/Domain: fibronectin type II repeat homology <2F2>

F:470-508/Domain: fibronectin type I repeat homology <1F7>

F:518-555/Domain: fibronectin type I repeat homology <1F8>

F:561-599/Domain: fibronectin type I repeat homology <1F9>

Alignment Scores:

Pred. No.: 2, 02e-35 Length: 2386

Score: 640.50 Matches: 208

Percent Similarity: 36.79% Conservative: 69

Best Local Similarity: 27.62% Mismatches: 153

Query Match: 17.19% Indels: 324

DB: 1 Gaps: 28

US-09-940-235-12 (1-2096) x FNUU (1-2386)

OY 5 GACCATCATGTTGTTGCTCAGGTCGACAGAGCTTTGACAGACAGAGTCGCTTACAGT--- 61

Db 122 AapCysThrCysIleGlyAlaGlyArgGlyArgIleSerCysThrIleAlaAsnArgCys 141

OY 62 -----TCGTCGCGGTATCGGATTCGATTCGCTTAACGACGTAAGGACACC 106

Db 142 HisGluGlyGlyGlnSerThr-LysIleGlyAspThrTrp-----ArgArgArg 157

OY 107 CCGCAGCCTTAAGCCGGGTCTCTCAACGACAGGACGACATGTCGACCCCGTGGCCAGGA 166

Db 157 OhIeGluThr----- 160

OY 167 CCCAAGCGTCCCGGAATCTCGATCCCGCCGAAATTAAATACGACTCACTATAGGAGACCA 226

Db 160 ----- 160

[illegible]

QY 1159 GTGACACATCATCTCAAGATTACTGCTCAGACCAAGACATTTAAACAAAACC 1218
||| ||| |||
Db 407 -----LeuValGlnThrGlnGly----- 412
QY 1219 ACCGAGGCTATGCAATTTATGACAGCGATCCGTCATGCTCACTGACATGATGATTT 1278
||| ||| |||
Db 413 -----GlyAsnSerAsnGlnValAlaLeuGlyHis----- 421
QY 1279 TCCGTCAGATTTTACCAATGATGACAGAGTTACTTACCGCTGTTAAATCGGGAACAAG 1338
||| ||| |||
Db 422 -----PheProheLeuTyrAsnAsnHis----- 429
QY 1339 CTATAGATGATCAATAAATAATGCTGTGAATGAGAAATAAACAACACTGACCTGATCT 1398
||| ||| |||
Db 430 -----AsnTyrThrAspCysThrS----- 436
QY 1399 CTGGAATAATTTACGCTCTTAAATAAGGGAAAAGCCGATATGATCCCTTGTGATCCGATC 1458
||| ||| |||
Db 436 ergLugly-----ArgArgAspA 442
QY 1459 ACTTGAAA---CTGTTACCATCAATACGTTGATGTCGATACCAAC-----G 1503
||| ||| |||
Db 442 smetLylTyrCysGlyThrThrGlnAsnTyrAspAlaAspGlnLysPheGlyPheCysP 462
QY 1504 AATGCTTAAAGTAGAGAGCTCTTAAACAGCTAGCGAAGCTAATTAGACTTCAGAGATT 1563
||| ||| |||
Db 462 rometAlaAlaHisGlnGluLeuLeuTyrThrThrAsnGlu----- 474
QY 1564 TATACGATTCCTGCTATAGGCTTAAAGCTTAACTTACCAACAATCTCGATGCTTTGGTATTA 1623
||| ||| |||
Db 475 -----GlyValM 477
QY 1624 TGGAATATACCTTAACTGGAATAAGTAGAGATATACGATGACACCAACCGTATATTA 1683
||| ||| |||
Db 477 et---TyrArgLLeuLysArgGlnTyrPAspLysGlnHisAspMetGlyHisSmetMetArgC 496
QY 1684 CCGTTTATATAGGAGCAAGCAAGCCGGAAGAGAAATGCTAGCTACCTTAACTGCTGAGTGG 1743
||| ||| |||
Db 496 ysthrCysValGly---AsnGlyArgGlyGlnTyrThrCysTyrAlaTyrSer----- 512
QY 1744 GCCAGGCGCAACAGTTGTATCCATAGCTAGAGAAAGTTTGTATGATCGCTGCGAGATT 1803
||| ||| |||
Db 513 -----GlnLeuArgAspGlnCysIle-----ValAspAspIleT 524
QY 1804 CCTATGTGTGTGGAGAAACGTGGAGAAAGCCCTACCAAGCGTGTGATGATGATTTGA 1863
||| ||| |||
Db 524 hrTyrAsnValAsnAspThrPheHisLysArgHisGlnGlnLysHisMetLeuAsnCysE 544
QY 1864 CTTCGCTGGAGAGAGCGAGCGGAGCATCTTGCACTTCTAAGTAATGATGCAAGATC 1923
||| ||| |||
Db 544 hrCysPheGlyGlnGlnLysArgLysTyrLysCysAspProValAspGlnCysGlnAspS 564
QY 1924 AGGACCAAGAGCATCTTATGAAATTTGAGACACTGAGCAAGAGAAATTAATCGAGAA 1983
||| ||| |||
Db 564 ergLutThrGlyThrPheTyrGlnLysLysAspSerTrpGluLysTyrValHis---GlyV 583
QY 1984 ACCTGCTCAGTGCATCTTGACAGCAAGCGCGAGAGAGTGAAGTGTGAGAGGCACA 2043
||| ||| |||
Db 583 alaTyrGlnGlnCysTyrCysTyrGlnArgGlyLysGlnGlnTyrHisCysGln----- 600
QY 2044 CCTGTGACAGACCATCGGCGGATCTGGGCC 2078
||| ||| |||
Db 601 --ProLeuGlnThrTyrProSerSerGlyPro 611

RESULT 9

FNBO

fibronectin - bovine

C/Species: Bos primigenius taurus (cattle)

C/Date: 31-Dec-1988 #sequence revision 31-Dec-1988 #text_change 09-Jul-2004

C/Accession: A26452; A26453; A23292

R/Skorenskaard, K.; Jensen, M.S.; Sahl, P.; Petersen, T.B.; Magnusson, S.

Eur. J. Biochem. 161, 441-453, 1986

A/Title: Complete primary structure of bovine plasma fibronectin.

A/Reference number: A26452; MUID:87054047; PMID:3780752
A/Accession: A26452
A/Molecule type: protein
A/Residues: 1-2265 <SKO>
A/Cross-references: UNIPROT:P07589; UNIPARC:UPI000012A7BE
R/Kornblith, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
Proc. Natl. Acad. Sci. U.S.A. 80, 3218-3222, 1983
A/Title: Isolation and characterization of cDNA clones for human and bovine fibronectin
A/Reference number: A21165; MUID:83221567; PMID:6304699
A/Accession: B21165
A/Molecule type: mRNA
A/Residues: 2170-2265 <KOR>
A/Cross-references: UNIPARC:UPI000016C30C; GB:K00800; NID:G163055; PIDN:AAA30521.2; PID
R/Petersen, T.B.; Skorenskaard, H.C.; Skorenskaard, K.; Vibe-Pedersen, K.; Sahl, P.; Sott
Proc. Natl. Acad. Sci. U.S.A. 80, 137-141, 1983
A/Title: Partial primary structure of bovine plasma fibronectin: three types of interna
A/Reference number: A23292; MUID:83117805; PMID:6218503
A/Accession: A23292
A/Molecule type: protein
A/Residues: 1-16, 'C', 18-20, 'S', 22-432; 447-463; 1367-1517; 1567-1673; 2062-2176, 'N', 2178-22
A/Cross-references: UNIPARC:UPI000017433B; UNIPARC:UPI000017433F; UNIPARC:UPI0000174340
C/Comment: Cys-1201 and Cys-2015 have free sulfhydryl groups.
C/Comment: The plasma fibronectin molecule consists of two chains, which are connected
C/Comment: fibronectins bind cell surfaces and various compounds including collagen, fil
aling, and maintenance of cell shape.
C/Comment: Plasma fibronectin is synthesized by hepatocytes.
C/Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II rep
C/Keywords: acute phase; alternative splicing; collagen binding; duplication; extracell
F/21-241/Domain: fibrin and heparin binding <FBR>
F/21-56/Domain: fibronectin type I repeat homology <1F1>
F/110-148/Domain: fibronectin type I repeat homology <1F2>
F/155-194/Domain: fibronectin type I repeat homology <1F3>
F/200-239/Domain: fibronectin type I repeat homology <1F4>
F/217-577/Domain: collagen binding <CBR>
F/277-311/Domain: fibronectin type I repeat homology <1F5>
F/329-370/Domain: fibronectin type II repeat homology <1F6>
F/389-430/Domain: fibronectin type II repeat homology <2F2>
F/439-477/Domain: fibronectin type I repeat homology <1F7>
F/487-524/Domain: fibronectin type I repeat homology <1F8>
F/530-568/Domain: fibronectin type I repeat homology <1F9>
F/578-661/Domain: fibronectin type III repeat homology <FN3A>
F/688-770/Domain: fibronectin type III repeat homology <FN3B>
F/779-860/Domain: fibronectin type III repeat homology <FN3C>
F/875-957/Domain: fibronectin type III repeat homology <FN3D>
F/965-1054/Domain: fibronectin type III repeat homology <FN3E>
F/1055-1134/Domain: fibronectin type III repeat homology <FN3F>
F/1142-1227/Domain: fibronectin type III repeat homology <FN3G>
F/1235-1318/Domain: fibronectin type III repeat homology <FN3H>
F/1326-1404/Domain: fibronectin type III repeat homology <FN3I>
F/1410-1517/Domain: cell attachment <CAD>
F/1416-1502/Domain: fibronectin type III repeat homology <FN3J>
F/1493-1495/Region: cell attachment (R-G-D) motif
F/1510-1592/Domain: fibronectin type III repeat homology <FN3K>
F/1600-1870/Domain: heparin binding <HB2>
F/1600-1862/Domain: fibronectin type III repeat homology <FN3L>
F/1692-1773/Domain: fibronectin type III repeat homology <FN3M>
F/1781-1863/Domain: fibronectin type III repeat homology <FN3N>
F/1970-1972/Region: cell attachment (R-G-D) motif
F/1982-2062/Domain: fibronectin type III repeat homology <FN3O>
F/1985-2216/Domain: fibrin binding <FB2>
F/2085-2124/Domain: fibronectin type I repeat homology <1F10>
F/2130-2167/Domain: fibronectin type I repeat homology <1F11>
F/2174-2209/Domain: fibronectin type I repeat homology <1F12>
F/Modified site: pyroliidone carboxylic acid (Gln) #status experimental
F/3/Cross-link: isopeptide (Gln) (interchain to Lys N6-amino of fibrin) #status exp
F/21-47,45-56,66-94,92-104,110-138,136-148,155-184,182-194,200-229,227-239,277-304,302-3
7,2155-2167,2174-2200,2198-2209/Dissulfide bonds: #status predicted
F/1205,1692/Binding site: carbohydrate (Asn) (covalent) #status absent
F/1943,1944/Binding site: carbohydrate (Thr) (covalent) #status experimental
F/2246/Dissulfide bonds: interchain (to 2250) #status predicted
F/2250/Dissulfide bonds: interchain (to 2246) #status predicted

Db 492 leThrTyAsnValAsnAspThrPheHisLysArgHisGluGluGlnHisMetLeuAsn 512
 QY 1861 GTACTTGCCTGGAGAGAGCAGCAGCATCTTGCATCTTGTAGAAATAGATGCAAG 1920
 Db 512 ySThrCySPheIylGlnGlyArgGlyArgTrpLysSAspProValAspGlnCysGln 532
 QY 1921 ATCAGACCAAGAACATCTTATAGAAATGGAGACACTGGAGCAAGAAAGATTAATCAG 1980
 Db 532 spSerGluThrArgThrPheTyGlnIleGlyAspSerTrp--GluLysTyLeuGln 551
 QY 1981 GAACCTGCTCCAGGATCTGCACAGGCAACGCCGAGAGAGAGTGGAAGTGTAGAGGC 2040
 Db 551 lYValArgTyGlnCysTyTyGlyArgGlyIleGlyGluTrpAlaCysGln---- 569
 QY 2041 AACCTCTGTGACACCATGACGAGCGGATCTGGCCCC 2078
 Db 570 -----ProLeuGlnThrTyProAspThrSerGlyPro 580
 RESULT 10
 S77671
 Streptokinase A (EC 3.4.-.-) (alleles 2 and 3) - Streptococcus pyogenes (fragment)
 C:Species: Streptococcus pyogenes
 C>Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 22-Jun-1999
 C:Accession: S77671; S77672
 R:Kapur, V.; Kanjilal, S.; Hamrick, M.R.; Li, L.L.; Whitlam, T.S.; Sawyer, S.A.; Musser, M.I. Microbiol. 16, 509-519, 1995
 A:Title: Molecular population genetic analysis of the streptokinase gene of Streptococcus A:Reference number: S77671; MUID:96037795; PMID:756511
 A:Accession: S77671
 A>Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-128 <KAB>
 A:Cross-references: UNIPARC:UPI0000000562; EMBL:U25853; NID:g818908; PIDN:AAA85729.1; PI
 A:Experimental source: strain E71/M1
 A:Note: allele 2
 A:Accession: S77672
 A>Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-128 <KAB>
 A:Cross-references: UNIPARC:UPI0000000562; EMBL:U25854; NID:g818910; PIDN:AAA85730.1; PI
 A:Experimental source: strain E2/M3
 A:Note: allele 3
 C:Genetics:
 A:Gene: ska
 C:Superfamily: streptokinase
 C:Keywords: hydrolase; plasminogen activator; virulence
 Alignment Scores:
 Pred. No.: 3,14e-34 length: 128
 Score: 620.00 Matches: 118
 Percent Similarity: 95.31% Conservative: 4
 Best Local Similarity: 92.19% Mismatches: 6
 Query Match: 16.64% Indels: 0
 DB: 2 Gaps: 0
 US-09-940-235-12 (1-2096) x S77671 (1-128)
 QY 1008 GTGGCGGTAGACCATATAAAGAAAAACCAATACAAACCAAGCGAATCTGTGATGTC 1067
 Db 1 ValArgValArgProCtyLysGluLysProValGlnAsnGlnAlaLysSerValAspVal 20
 QY 1068 GAATATACTGTACAGTTTACTCCCTTAAACCTGATGACGATTTGACACCGAGTCTCAAA 1127
 Db 21 GluLysThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 40
 QY 1128 GATCTAGCTATTGAACAACCTAGCTATCGGTGACACATCATCTCAAGAAATACATA 1187
 Db 41 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 60
 QY 1188 GCTCAGACCAAGCATTTTAAACAAACACACCGAGCTATGATGATTTATGAAACGTGAC 1247
 Db 61 AlaGlnIleGlnSerIleLeuAsnLysThrHisProGlyTyThrIleTyGlnLysAsp 80

QY 1248 TCCTCAATCGCTACCTACATGACATGATTTTCCGTACAGATTTTACCAATGATGACAG 1307
 Db 81 SerSerIleValThrHisAspAsnAspIlePheHisGlnThrIleLeuProMetCAspGlnGlu 100
 QY 1308 TTATCTTACCGTGTAAATAATCGGACACAGCTTATAGATCATGATTAATAAATCGTGTG 1367
 Db 101 PheThrTyGlnHisValLysAsnArgGlnGlnAlaTyGlnIleAsnProLysThrGlyIle 120
 QY 1368 AATGAGAAATTAACAACATGAC 1391
 Db 121 LysGluLysThrAsnAsnThrAsp 128
 RESULT 11
 S14428
 fibronectin precursor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 09-Jul-2004
 C:Accession: S14428; S14455; A22319; S46203; S00459; A27252; I59049
 R:Hyne, R.O.
 Submitted to the EMBL Data Library, July 1989
 A:Reference number: S14428
 A:Accession: S14428
 A:Molecule type: mRNA
 A:Residues: 1-2477 <HYN>
 A:Cross-references: UNIPROT:P04937; UNIPARC:UPI000012ATC6; EMBL:X15906; NID:g56163; PID
 R:Schwarzbauer, J.E.; Patel, R.S.; Fonda, D.; Hyne, R.O.
 A:Title: Multiple sites of alternative splicing of the rat fibronectin gene transcript.
 A:Reference number: S12455; MUID:88054951; PMID:2445560
 A:Accession: S12455
 A>Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 609-1810,'T',1812-2283 <SCH>
 A:Cross-references: UNIPARC:UPI0000177AE; EMBL:X15906
 R:Tamkun, J.W.; Schwarzbauer, J.E.; Hyne, R.O.
 Proc. Natl. Acad. Sci. U.S.A. 81, 5140-5144, 1984
 A:Title: A single rat fibronectin gene generates three different mRNAs by alternative splicing.
 A:Reference number: A22319; MUID:84298097; PMID:6089177
 A:Accession: A22319
 A:Molecule type: DNA
 A:Residues: 2052-2237 <TAM>
 A:Cross-references: UNIPARC:UPI0000177AEB
 R:Falkenberg, C.; Englund, J.J.; Thogersen, I.B.; Salvesen, G.; Akerstrom, B.
 Biochem. J. 301, 745-751, 1994
 A:Title: Isolation and characterization of fibronectin-alpha(1)-microglobulin complex in A:Reference number: S46203; MUID:94330948; PMID:7519849
 A:Accession: S46203
 A>Status: preliminary
 A:Molecule type: protein
 A:Residues: 1183-1192; GLN',1268,'P',1270-1271,'D',1273,'CF',1276,'PY',1385-1399 <FAL>
 A:Cross-references: UNIPARC:UPI0000177AEC; UNIPARC:UPI0000177AED; UNIPARC:UPI0000177AEE
 R:Patel, R.S.; Odehmat, E.; Schwarzbauer, J.E.; Hyne, R.O.
 EMO J. 6, 2565-2572, 1987
 A:Title: Organization of the fibronectin gene provides evidence for exon shuffling during A:Reference number: S00459; MUID:88054950; PMID:3119323
 A:Accession: S00459
 A:Molecule type: DNA
 A:Residues: 1-139;2382-2477 <PAT>
 A:Cross-references: UNIPARC:UPI000017095C; UNIPARC:UPI0000177AEF; EMBL:X05831
 R:Schwarzbauer, J.E.; Tamkun, J.W.; Lemischka, I.R.; Hyne, R.O.
 Cell 33, 421-431, 1983
 A:Title: Three different fibronectin mRNAs arise by alternative splicing within the codi A:Reference number: A27252; MUID:84082067; PMID:6317187
 A:Accession: A27252
 A:Molecule type: mRNA
 A:Residues: 1586-1720,'T',1722,1813-2477 <SC2>
 A:Cross-references: UNIPARC:UPI0000177AF0
 R:Odehmat, E.; Tamkun, J.W.; Hyne, R.O.
 Proc. Natl. Acad. Sci. U.S.A. 82, 6571-6575, 1985
 A:Title: Repeating modular structure of the fibronectin gene: Relationship to protein str A:Reference number: I59049; MUID:86016741; PMID:3863113
 A:Accession: I59049


```

Qy 1186 TAGCTCAAGCACAAGCATTTTAAACAAAACCCAGGCTTATGATTTATGAACGTG 1245
Db ||| |||
407 euValGlnThrArgely-----GlyAsn 415
Qy 1246 ACTCTCAATGCTCATGATGACATTTTCGTACGATTTTACCAATGATCAAG 1305
Db ||| |||
415 efaSnglYAlaIeuCysHis-----421
Qy 1306 AGTTTACTTACCGTGTAAAAATCGGAAACAAGCTTATAGATCAATAAAAATCTGTC 1365
Db ||| |||
422 --PheProPheIeuYrSerAenArg-----429
Qy 1366 TGAATGAAGAATAAACAACACTGACCTGATCTTGAAGAAATTAAGTCTTAAAAAG 1425
Db ||| |||
430 -----AsnYrSerAenPheCysThrSerGlnY-----438
Qy 1426 GGGAAAGCGGTATGATCCCTTTGATTCGACGATCACTTGA---CTGTTCACCATCAAT 1482
Db ||| |||
439 -----ArgArgAspAsnMetLysTrpCysGlyThrThrGlnA 451
Qy 1483 AGTTGATGTCGATACCAAC-----GAATTGTAAGAAAGTGAAGCAGCTCTTAA 1530
Db ||| |||
451 snYrAspAlaAspGlnLysPheGlyPheCysProMetAlaAlaHisGlnGlnIleCysT 471
Qy 1531 CAGCTAGCGAAGCATTAAGCTTGAAGATTTATGATTCCTCGTATGAGGCTTAAAC 1590
Db ||| |||
471 hrThrAsnGln-----474
Qy 1591 TACTCTACAACAATCTCGATGCTTTTGGTATATGACATTTACCTTAACTGAAAAATGAG 1650
Db ||| |||
475 -----GlyValMet---TyrArgIleGlyAspGlnTrpA 485
Qy 1651 AAGATATCAGATGACACCAACCGATCATTAACGTTTATAGGCAAGCAGCCGAAG 1710
Db ||| |||
485 sPlySglnHisAsp-----LeuG 491
Qy 1711 GAGAGATGCTAGTACATTAAGTGTGTGTGCGCAGCGCA---CAGATTGTACCC- 1766
Db ||| |||
491 LysHisMetMetArgCysThrCysValGlyAsnGlyArgGlyIleTrpAlaCysIleProT 511
Qy 1767 -----ATAGCTGAGAAAGTTTGTATGATCATGCTGTGAGACTTCTTATGTGTGGAG 1818
Db ||| |||
511 yrSerGlnLeuArgAspGlnCysIle-----ValAspAspIleThrYrAsnValAsnA 529
Qy 1819 AAAGTGGGAGAGCCCTACCAAGCGTGTGATGATGATGATGATGATGATGATGATGATG 1878
Db ||| |||
529 sPTrPheHisIleYrGlnHisGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 549
Qy 1879 GCAGCGAGCGCATCTTCACTTCTAGAAATGATGATGATGATGATGATGATGATGATGAT 1938
Db ||| |||
549 LysArgGlyArgTrpLysCysAspProIleAspArgCysGlnAspSerGlnThrArgThr 569
Qy 1939 CTTATGATTTGAGACACTTGTGACCAAGAGATGATGATGATGATGATGATGATGATGATG 1998
Db ||| |||
569 heYrGlnIleGlyAspSerTrp---GlnLysPheValHisGlyValArgTrpGlnCysT 588
Qy 1999 TCTGCAAGCGCAAGCGCGGAGAGAGTGAAGTGTGAGAGGCAACCTTGTGACAGCA 2058
Db ||| |||
588 yrCysYrGlyArgGlyIleGlyIleGlyIleGlyIleGlyIleGlyIleGlyIleGlyIle 605
Qy 2059 CATGAGAGGATCTGCGCC 2078
Db ||| |||
605 yrProGlyThrThrIlePro 611

```

```

A:Reference number: S77671; MUID:96037795; PMID:7565111
A:Accession: S77680
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-128 <RAP>
A:Cross-references: UNIPROT:Q54687; UNIPARC:UP100000BD2D3; EMBL:U25862; NID:g818926; PIR:
A:Experimental source: strain ET51/M17
C:Genetics:
A:Gene: ska
A:Superfamily: streptokinase
C:Keywords: hydrolase; plasminogen activator; virulence

Alignment Scores:
Pred. No.: 8,68e-11 Length: 128
Score: 569.00 Matches: 110
Percent Similarity: 91.41% Conservative: 7
Best Local Similarity: 85.94% Mismatch: 11
Query Match: 15.27% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-12 (1-2096) x S77680 (1-128)

Qy 1008 GTGGCGGTTAGACCAATTAAGAAAAACAATACAAAACAGCGAAATCTGTGATGTG 1067
Db ||| |||
1 ValArgValArgProTrpLysGlnLysProIleGlnThrProAlaLysSerValAspVal 20
Qy 1068 GAATATATCTGTACAGTTACTTCCCTTAAACCTGTATGACGATTTTCAAGCCAGGCTCAAA 1127
Db ||| |||
21 ArgGlyThrValGlnPheThrProLeuAspProAspAspPheArgProValLys 40
Qy 1128 GATATCAAGCTATTTGAAAAACATAGCTATCGGTGACACACATGATGATGATGATGATGAT 1187
Db ||| |||
41 AsnThrIleLeuLeuLysThrIleAlaIleGlyGlyThrValHisSerGlnIleLeu 60
Qy 1188 GCTCAAGCACAAGCATTTTAAACAAAACCCAGGCTTATGATTTATGACGTGAC 1247
Db ||| |||
61 AlaGlnAlaGlnSerIleLeuAsnGlnSerHisProAspTrpThrIleYrGlnLysAsp 80
Qy 1248 TCTCAATGCTCACTGATGACATGATGATTTTCCGTACGATTTTACCAATGATCAAGAG 1307
Db ||| |||
81 SerSerIleValHisHisAspAspIlePheArgThrIleLeuProMetAspGlnGln 100
Qy 1308 TTTATCTACCGGTTAAAAATCGGGAACAGCTTATGATGATGATGATGATGATGATGATG 1367
Db ||| |||
101 PheThrThrHisIleLysAspArgGlnGlnAlaIleGlyIleAsnLysSerGln 120
Qy 1368 AATGAAGAATAAACAACACTGAC 1391
Db ||| |||
121 GlnGlnLysThrAsnThrAsp 128

RESULT 13
S77680
streptokinase A (EC 3.4.-.-) (allele 19) - Streptococcus pyogenes (fragment)
C:Species: Streptococcus pyogenes
C>Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004
C:Accession: S77688
R:Kapur, V.; Kanjilal, S.; Hamrick, M.R.; Li, L.L.; Whittam, T.S.; Sawyer, S.A.; Musser,
Mol. Microbiol. 16, 509-519, 1995
A:Title: Molecular population genetic analysis of the streptokinase gene of Streptococcus
A:Reference number: S77671; MUID:96037795; PMID:7565111
A:Accession: S77688
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-128 <RAP>
A:Cross-references: UNIPROT:Q54695; UNIPARC:UP100000BD8DC; EMBL:U25870; NID:g818942; PIR:
A:Experimental source: strain ET76/M72
C:Genetics:
A:Gene: ska
A:Superfamily: streptokinase
C:Keywords: hydrolase; plasminogen activator; virulence

Alignment Scores:
Pred. No.: 5.61e-30 Length: 128

```

```

Score: 557.00 Matches: 107
Percent Similarity: 89.84% Conservative: 8
Best Local Similarity: 83.59% Mismatches: 13
Query Match: 14.95% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-12 (1-2096) x S77688 (1-128)

QY 1008 GTGGCGCTTAAAGCATATTAAGAAAAACCAATACAAACCAAGCAATCTGTGATGNG 1067
Db 1 ValAGValAGPrCTrYrLysGluYsProlIeGlnThrProAlaYsSerValaPtle 20
QY 1068 GAATATACGTACAGTTTACTCCCTTAAACCTGATGACGATTTACAGACGGCTCAAA 1127
Db 21 ArgTYrThrValGlnPheThrProLeuAsnProAdaPAspPheYsProValLeuYs 40
QY 1128 GATCTAAGCTATTTAAACACTAGCTATCGGTGACACATCAATCTCAAGATTACTA 1187
Db 41 AspThrYsLeuLeuYsThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 60
QY 1188 GCTCAGACCAAGCATTTTAAACAAAACCAACGAGGTATACGATTTATGACGTGAC 1247
Db 61 AlaGlnAlaGlnSerIleLeuIleGluSerHisProAspTYrThrIleTYrGluArgAsp 80
QY 1248 TCCTCAATCGTCACTCATGACATGACATTTTCCTGATGATTTTACCAATGATCAAGAG 1307
Db 81 SerSerIleValThnHisAspAsnAspIlePhehAlyGlnIleLeuProThrAspGln 100
QY 1308 TTTAACCTACCGTGTAAAAATCGGAGAACGCTTATAGATCAATAAAAATCTGCTG 1367
Db 101 PheThrYrHisValIysAlaAsnArgGluGlnAlaTYrYrYAlaAsnSerYsThrAspIle 120
QY 1368 AATGAAGAATTAACAACATGAC 1391
Db 121 LysGluYsThrAsnAsnThrAsp 128

RESULT 14
A43908
fibronectin - African clawed frog
C/Species: Xenopus laevis (African clawed frog)
C/Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 12-Feb-1999
C/Accession: A43908
R/DeSimone, D.W.; Norton, P.A.; Hynes, R.O.
Dev. Biol. 149, 357-369, 1992
A/Title: Identification and characterization of alternatively spliced fibronectin mRNAs
A/Reference number: A43908; MUID:92111942; PMID:1730390
A/Accession: A43908
A/Status: nucleic acid sequence not shown; not compared with conceptual translation
A/Molecule type: mRNA
A/Residues: 1-2481 <DBS>
A/Cross-References: UNIPARC:UPI0000177AE7; GB:M77820
A/Note: sequence extracted from NCBI Backbone (NCBI:77473)
C/Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II repe
C/Keywords: duplication; extracellular matrix; glycoprotein; heterodimer
F/55-90/Domain: fibronectin type I repeat homology <1F1>
F/100-138/Domain: fibronectin type I repeat homology <1F2>
F/144-182/Domain: fibronectin type I repeat homology <1F3>
F/189-228/Domain: fibronectin type I repeat homology <1F4>
F/234-273/Domain: fibronectin type I repeat homology <1F5>
F/309-343/Domain: fibronectin type I repeat homology <1F6>
F/361-402/Domain: fibronectin type II repeat homology <2F1>
F/421-462/Domain: fibronectin type II repeat homology <2F2>
F/471-509/Domain: fibronectin type I repeat homology <1F7>
F/519-556/Domain: fibronectin type I repeat homology <1F8>
F/562-600/Domain: fibronectin type I repeat homology <1F9>
F/610-693/Domain: fibronectin type III repeat homology <FN3A>
F/719-801/Domain: fibronectin type III repeat homology <FN3B>
F/810-881/Domain: fibronectin type III repeat homology <FN3C>
F/906-988/Domain: fibronectin type III repeat homology <FN3D>
F/996-1077/Domain: fibronectin type III repeat homology <FN3E>
F/1086-1165/Domain: fibronectin type III repeat homology <FN3F>
F/1173-1258/Domain: fibronectin type III repeat homology <FN3G>
F/1266-1349/Domain: fibronectin type III repeat homology <FN3H>

```

```

F/1357-1440/Domain: fibronectin type III repeat homology <FN3I>
F/1448-1530/Domain: fibronectin type III repeat homology <FN3J>
F/1538-1620/Domain: fibronectin type III repeat homology <FN3K>
F/1615-1617/Region: cell attachment (R-G-D) motif
F/1632-1714/Domain: fibronectin type III repeat homology <FN3L>
F/1722-1804/Domain: fibronectin type III repeat homology <FN3M>
F/1812-1894/Domain: fibronectin type III repeat homology <FN3N>
F/1904-1985/Domain: fibronectin type III repeat homology <FN3O>
F/1993-2075/Domain: fibronectin type III repeat homology <FN3P>
F/2197-2277/Domain: fibronectin type III repeat homology <FN3Q>
F/2301-2340/Domain: fibronectin type I repeat homology <1F10>
F/2346-2383/Domain: fibronectin type I repeat homology <1F11>
F/2390-2425/Domain: fibronectin type I repeat homology <1F12>
F/55-81-79-90, 100-128, 126-138, 144-172, 170-182, 189-218, 216-228, 234-263, 261-273, 309-336, 337-373, 2371-2383, 2390-2416, 2414-2425/Disulfide bonds: #status predicted
F/2459/Disulfide bonds: Interchain (to 2463) #status predicted
F/2463/Disulfide bonds: Interchain (to 2459) #status predicted

Alignment Scores:
Pred. No.: 1,29e-29 Length: 2481
Score: 554.50 Matches: 185
Percent Similarity: 36.56% Conservative: 91
Best Local Similarity: 24.50% Mismatches: 150
Query Match: 14.88% Indels: 330
DB: 2 Gaps: 28

US-09-940-235-12 (1-2096) x A43908 (1-2481)

QY 5 GACCATTCATGTTGTGCTCAGGTGCGACAGCTTTTGACAGACAGTGGCTTCACGT--- 61
Db 125 AspCYrThrCYsrlleGlyAlaGlyArgGlyArgIleSerCYsThrIleAlaAsnArgCYs 144
QY 62 -----TCGCTCGGCTATCGGTGATTCATTTCTGCTAACCAAGGCAACC 106
Db 145 HisGluGlyGlyGlnSerTYr-LysrlleGlyAspThrTrp-----ArgArgPr 160
QY 107 CCGCAGCCTAGCCGGGTCTCAACAGACAGACAGATGATGCGACCCGTGGCCAGCA 166
Db 160 ohlsGluThr----- 163
QY 167 CCCAAGCTGCCCGAGATCTCGATCCCGGAAATTATACGACTCATATAGGAGACCA 226
Db 163 ----- 163
QY 227 CAACGGTTCCCTTAGAA---ATAATTGTTTAACTTAAAGAGATATATACATGG 283
Db 164 -GlyGlyTYrMetLeuGlyCYsValCYsLeuGlyAsnGlyGlyu- TrpThr---- 181
QY 284 TGAACACACACAGATTTGATCCATGACGAGAAAGTGTGATCATGCTGCGGACTT 343
Db 182 -----CYsLysProValAlaGluArgCYsTYrAspAsnThrAlaGlyThrs 197
QY 344 CCTATGTGTCGGAAGAACGTGGGA----- 368
Db 197 erTYrValValGlyGlnThrTrpGluYsProTYrGlnGlyTYrPmetMetValAspCYsT 217
QY 369 -----GAAGCAGGAGCAGCATCATCTTGACCTTGAAGATATGATGACATC 417
Db 217 hrcYsLeuGlyGlnGlyAsnGlyArgIleThrcYsSerSerYsAsnArgCYsAsnAspG 237
QY 418 AGGACACAGGACATCTTATGAATTGAGACACCTGAGCAAGAGATATATGAGAA 477
Db 237 lnsPThrYsThrSerTYrArgIleGlyAspThrTrpSerYsThrAspThrArgGlyA 257
QY 478 ACCGTGTCAGGTCATCTGCAACAGGCAAGGCGGAGAGAGTGAAGTGAAGGACCA 537
Db 257 snLeuLeuGlnCYsrlleCYsThnGlyAsnGlyArgGlyGlnTrpYsCYsGluArgHis 277
QY 538 CCTGTGGAACACATGAGGAGATCGGCCCTTACACCATGTTCTGATTCGTGAC 597
Db 277 eserAlaGlnAlaThrGlyThnGlySerAsnProIleThraAnIleGlnThrAla---- 295
QY 598 CTGAGTGGCTGCTAGACCGTCCATCTGTCAACACACGCAATTTGTTTGAAGCTTGG 657

```



```

Db      296  -----|||:::|||||-----AspSerGlnLeu-----303
Oy      658  GTACTGTGAGGGAGCAATCAAGACATAGTCTTAATTTTGAATGATCAATACAT717
Db      303  -----303
Oy      718  CACGACCTGCTATGAGAGAAAGACAGAGCGTTAAGTCCAAATCAAAACATTTG777
Db      304  -----GluProtyrG307
Oy      778  CTACTGATAGTGCGCGGATGTCACTAAACTTGAAGAAAGCTGACTTACTAAAGCTATTTC837
Db      307  1y-----307
Oy      838  AAGAACATTTGATCGCTAACGTCACAGTAACGACGACTTGAAGTATGATTGTTG897
Db      307  -----307
Oy      898  CAAGCGATGCAACATTAAGTATGATGCAACGCAAGGTCTATTGCTGACAAAGATGTT957
Db      308  -----HisCysValThrAsp---AsnGlyValLeuTyrSerLeu-----GlyM321
Oy      958  CGGTAACTTGCCCGACCACTGTCCAGAAATTTTGTAAAGCGGACATGCGCGGTTA1017
Db      321  eCArgTTrpLeuArgThrGlnGlySerTyrGlnMetLeuCysThr-----335
Oy      1018  GACCATATTAAGAAAAACCAATACAAACCAAGCGAAATCTGTGATGTGGAATATATCTG1077
Db      335  -----335
Oy      1078  TACAGTTACTTCCTTAAACCTGATGACGATTTCAAGACGAGTCCAAAGATCACTAAGC1137
Db      335  -----335
Oy      1138  TATTGAAAACTAGTACTGATGATGACATCAATCTCAAGATTAAGTACTCAAGCAC1197
Db      336  -----CysLeuGlyAsnGlyValSerCysGlnGluThrValAlaThrIleT351
Oy      1198  AAGACATTTTAAACAAAAAC-----CACCCAGGCTATA1230
Db      351  hPheGlyGlyAsnAlaAsnGlyGluProCysAlaIleProPheThrHisAspGlyLysT371
Oy      1231  CGATTATGACGTGACTCCTCA-----ATGTCATCTC1263
Db      371  hTyrTyrSerCysThrSerGlnGlyArgGlnAspGlyLysLeuTyrCysAlaThrThrS391
Oy      1264  ATGACATGACATTTTCCGTACGATTTTACCAATGATCAAGATTTACTTACTCCGTGTTA1323
Db      391  eAsnTyrAsp-----SerAspLysLysTyrSerPheCysThr-----403
Oy      1334  AAAATCGGACAGAGCTTATAGATCAATAAAAATCTGTT-----1364
Db      404  -----GluGlnLeuAlaLeuValGlnThrArgGlyGlyAsnSerAsnGlyAlaLeuG421
Oy      1365  -----CTGATGAAGAATAAACAACACTGACCTGATCTCGAAGAAATATT1410
Db      421  yAsnProPheProPheLeuTyrAsnAsnLeuAsnTyrThrAspCysThrSerGln-----438
Oy      1411  ACGTCTTAAAAAAGGGGAAAGCCGTATGATCCCTTTGATGCGACGACTTGAAATGTT1470
Db      439  -----GlyArgGlnAsp-----442
Oy      1471  TCACATCAATATACGTTGATGTCGATCAAC-----1502
Db      443  -----SerMetLysTyrCysGlyThrThrAlaSerTyrAspAlaAspGlnLysPheGlyPheC462
Oy      1503  --GAATGTCTAAAAAGTGAAGCAAGCTCTTAACAGCTTGAAGCAAGTAACTTAAGCTTACAGAG1560
Db      462  ySPrometLalaIariGlnGlnGlnTyrSerThrThrAsnGln-----475
Oy      1561  ATTATAGATCTCTGATGATAGGCTAAACTACTACTCAACAACATTCGATGCTTTGGTA1620

```

```

Db      476  -----GlyV477
Oy      1621  TTATGACTATACCTTAACGTGAAAAGTAGAGATTAATACAGATGACACCAAGCTATCA1680
Db      477  alMet---TyrArgValGlyAspGlnTyrAspLysGlnHisAspGlnGlyHisMetCys496
Oy      1681  TAACTGTTATATAGGCGAAG---CGACCCGAGAGAGAGAAATGCTAGCAATTTAAGCTG1737
Db      496  rgCysThrCysValGlyAsnGlyArgGlyGluTyrSerCysValAlaTyrSer-----513
Oy      1738  GTGATGCGCAGCGCCACACATTTGATACCATAGCTGAGAAAGTGTTTTATCATGCTGCTG1797
Db      514  -----GlnLeuLysAspGlnCysIle-----ValAspG523
Oy      1798  GAATCTCTATGTGTGCGAGAAAAGTAGGAGAAAGCCCTACCAAGCGCTGATGATGCTAG1857
Db      523  LysLeuThrTyrAsnValAsnSerSerPheThrLysLeuHisGlnGlnGlyHisMetCys543
Oy      1858  ATTGTACTTGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG1917
Db      543  snCysThrCysPheGlyGlnGlyArgGlyArgTyrCysAspAlaIleAspGlnCysG563
Oy      1918  ACGATCAGACACAAAGACATCTATAGAAATTTGAGACACTGAGACAAAGAGATATATC1977
Db      563  LnaPthrGlnThrArgGlnPheThrGlnIleGlyAspSerTrp---GluLysHisLeuG582
Oy      1978  GAGGAACCTGCTCCAGTGCATCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG2037
Db      582  LngLysValGlnTyrGlnCysTyrCysTyrGlyLysGlyIleGlyGlnTyrHisCysGln601
Oy      2038  GGCACACCTGCTGACAGACATGACAGCGAGCTGAGCCCTG2078
Db      602  -----ProLeuSerThrSerGlnAlaGlyThrGlyPro612

```

```

RESULT 15
S77679
Streptokinase A (EC 3.4.-.-) (allele 10) - Streptococcus pyogenes (fragment)
C:Species: Streptococcus pyogenes
C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004
C:Accession: S77679
R:Kapur, V.; Kanjilal, S.; Hamrick, M.R.; Li, L.L.; Whitlam, T.S.; Sawyer, S.A.; Musser
Mol. Microbiol. 16, 509-519, 1995
A:Title: Molecular population genetic analysis of the streptokinase gene of Streptococcus
A:Reference number: S77671; MUID:56037795; PMID:7565111
A:Accession: S77679
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-128 <KAP>
A:Cross-references: UNIPROT:Q54686; UNIPARC:UPI000008BDF6; EMBL:U25861; NID:9818924; PIR
A:Experimental source: strain ET50/M43
C:Genetics:
A:Gene: sika
C:Superfamily: streptokinase
C:Keywords: hydrolase; plasminogen activator; virulence

Alignment Scores:
Pred. No.: 1,22e-29 Length: 128
Score: 552.00 Matches: 106
Percent Similarity: 89.06% Conservative: 8
Best local Similarity: 82.81% Mismatches: 14
Query Match: 14.81% Indels: 0
DB: 2 Gaps: 0

```

```

US-09-940-235-12 (1-2096) x S77679 (1-128)
Oy      1008  GTGCGCGTGTGACCATTAAGAAAGCAATACAAACCAAGGAATCGTTGATGTG1067
Db      1  ValArgValArgProTyrLysGlnLysProIleGlnThrProAlaLysSerValAspIle20
Oy      1068  GAATATATCTGTATACGTTACTTCCCTTAAACCTGATGACGATTTCAAGACAGTCTCAAA1127
Db      21  ArgTyrLysValGlnPheThrProLeuAsnProAspAspPheThrProValLeuLys40

```



```

Qy 1128 GATACCTAAGCTATTGAAAACTAGCTATCGGTGACACCATCAATCTCAAGATTACTA 1187
    |||
Db 41 AsptHrIysleuLeuIysThrIeuAlaIeGIyAspThrIleThrSerGIingIuLeu 60
    |||
Qy 1188 GCTCAAGCACAAAAGCATTTTAAACAAAACCAAGGCTATACGATTATGAAAGGTGAC 1247
    |||
Db 61 AlaGIAlaGIInSerIleIeuIleGIInSerHISProAsnTyThrIleHISGIuArGAsp 80
    |||
Qy 1248 TCCGCAATCGTCACATGACAAATGACATTTCCGTACGATTTTACCAATGATCAAGAG 1307
    |||
Db 81 SerSerIleValThrHisaspAsnAspIlePheArGThrIleIeuProThrAspGIingIu 100
    |||
Qy 1308 TTTACTTACCGTGTAAATAATCGGGAACAAGCTTATAGGATCAATAATAAATCGGTG 1367
    |||
Db 101 PheThrTyHisValIysAsnArGIingIuAlaTyTyIysAlaAsnSerIysTingIyIle 120
    |||
Qy 1368 AATGAAGAATTAACACACTGAC 1391
    |||
Db 121 LysGIuLysThrAsnAsnThrAsp 128
    |||

```

Search completed: January 28, 2006, 02:38:18
 Job time : 56.1209 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 28, 2006, 01:48:09 ; Search time 225.882 Seconds
(without alignments)
13093.422 Million cell updates/sec

Title: US-09-940-235-12

Perfect score: 3726

Sequence: 1 cgaagaccatcatgtgtgtt.....ccttcacgatgtcgttag 2096

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2166443 segs, 705528306 residues

Total number of hits satisfying chosen parameters: 4332886

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame_n2p.model -DEV=xlp
-O=/cgn2.1/USPTO_spool_p/US09940235/runat_27012006_144218_27563/app_query.fasta.1.7708
-DB=UniProt -QFMT=fastan -SUFFIX=rup -MINMATCH=0.1 -LOOPEXT=0 -LIST=45
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NOM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09940235 @CGN_1.1.1355 @runat_27012006_144218_27563 -NCPU=6 -ICPU=3
-NO MMAP -LARGEUSER -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : UniProt_05.80.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1984	53.2	440	1 STRP_STRREQ	P00779 streptococc
2	1985	52.5	414	2 Q53284 STRREQ	Q53284 streptococc
3	1943	52.1	436	2 Q6UK57 STRREQ	Q6UK57 streptococc
4	1942	52.1	440	1 STRP_STRS1	P10519 streptococc
5	1814	48.7	440	2 Q7X0Y1 STRPY	Q7X0Y1 streptococc
6	1802	48.4	440	2 Q7X0Y8 STRPY	Q7X0Y8 streptococc
7	1802	48.4	440	2 Q8K5R8 STRP3	Q8K5R8 streptococc
8	1794	48.1	440	1 STRP_STRPY	P10520 streptococc
9	1794	48.1	440	2 Q532X6 STRPY	Q532X6 streptococc
10	1787	48.0	440	2 Q7X0Y2 STRPY	Q7X0Y2 streptococc
11	1758	47.2	440	2 Q7X0Y7 STRPY	Q7X0Y7 streptococc
12	1758	47.2	440	2 Q5X9T6 STRP6	Q5X9T6 streptococc
13	1755	47.1	440	2 Q8NZ46 STRP8	Q8NZ46 streptococc
14	1723	46.2	440	2 Q7X0Y3 STRPY	Q7X0Y3 streptococc
15	1715	46.0	440	2 Q7X0Y0 STRPY	Q7X0Y0 streptococc
16	1715	46.0	440	2 Q7X0Y5 STRPY	Q7X0Y5 streptococc

17	1708	45.8	440	2 Q7X0X8 STRPY	Q7X0X8 streptococc
18	1686	45.2	440	2 Q7X0X9 STRPY	Q7X0X9 streptococc
19	1677	45.0	440	2 Q7X0X7 STRPY	Q7X0X7 streptococc
20	1670	44.8	440	2 Q57391 STRPY	Q57391 streptococc
21	1657	44.5	440	2 Q7X0Y6 STRPY	Q7X0Y6 streptococc
22	1652	44.3	440	1 STRQ_STRPY	Q7X0Y6 streptococc
23	1624	43.6	432	2 Q7X0Y4 STRPY	Q7X0Y4 streptococc
24	1624	43.6	432	2 Q7X0X2 STRREQ	Q7X0X2 streptococc
25	1624	43.6	432	2 Q7X0X3 STRREQ	Q7X0X3 streptococc
26	1624	43.6	432	2 Q7X0X4 STRREQ	Q7X0X4 streptococc
27	1624	43.6	432	2 Q7X0X5 STRREQ	Q7X0X5 streptococc
28	1624	43.6	432	2 Q7X0X6 STRREQ	Q7X0X6 streptococc
29	1624	43.6	432	2 Q7X0X7 STRREQ	Q7X0X7 streptococc
30	1624	43.6	432	2 Q7X0X8 STRREQ	Q7X0X8 streptococc
31	1624	43.6	432	2 Q7X0X9 STRREQ	Q7X0X9 streptococc
32	1624	43.6	432	2 Q7X0X0 STRREQ	Q7X0X0 streptococc
33	1624	43.6	432	2 Q7X0X1 STRREQ	Q7X0X1 streptococc
34	1624	43.6	432	2 Q7X0X2 STRREQ	Q7X0X2 streptococc
35	1624	43.6	432	2 Q7X0X3 STRREQ	Q7X0X3 streptococc
36	1624	43.6	432	2 Q7X0X4 STRREQ	Q7X0X4 streptococc
37	1624	43.6	432	2 Q7X0X5 STRREQ	Q7X0X5 streptococc
38	1624	43.6	432	2 Q7X0X6 STRREQ	Q7X0X6 streptococc
39	1624	43.6	432	2 Q7X0X7 STRREQ	Q7X0X7 streptococc
40	1624	43.6	432	2 Q7X0X8 STRREQ	Q7X0X8 streptococc
41	1624	43.6	432	2 Q7X0X9 STRREQ	Q7X0X9 streptococc
42	1624	43.6	432	2 Q7X0X0 STRREQ	Q7X0X0 streptococc
43	1624	43.6	432	2 Q7X0X1 STRREQ	Q7X0X1 streptococc
44	1624	43.6	432	2 Q7X0X2 STRREQ	Q7X0X2 streptococc
45	1624	43.6	432	2 Q7X0X3 STRREQ	Q7X0X3 streptococc

ALIGNMENTS

RESULT 1
STRP_STRREQ STANDARD; PRT; 440 AA.
AC P00779;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Streptococcus C precursor.
GN Name=ekc;
OS Streptococcus equisimilis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=119602;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=H46A;
RX MEDLINE=85232082; PubMed=2989113; DOI=10.1016/0378-1119(85)90145-3;
RA Makle H., Roe B.A., Ferretti J.U.;
RT "Nucleotide sequence of the streptokinase gene from Streptococcus equisimilis H46A.";
RL Gene 34:357-362(1985).
RN [2]
RP PROTEIN SEQUENCE OF 27-440.
RX MEDLINE=83127125; PubMed=6760891;
RA Jackson K.W., Tang J.;
RT "Complete amino acid sequence of streptokinase and its homology with serine proteases.";
RL Biochemistry 21:6620-6625(1982).
-!- FUNCTION: This protein is not a protease, but it activates plasminogen by complexing with it. As a potential virulence factor, it is thought to prevent the formation of effective fibrin barriers around the site of infection, thereby contributing to the invasiveness of the cells.
-!- Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

DR EMBL; K02986; AAA26974.1; -; Genomic DNA.
 DR EMBL; X72832; CAA51351.1; -; Genomic DNA.
 DR PIR; A00967; BZSO.
 DR PIR; A22801; A22801.
 DR PDB; 1BML; X-ray; C/D=38-399.
 DR PDB; 1L4D; X-ray; B=40-173.
 DR PDB; 1L4Z; X-ray; B=27-173.
 DR PDB; 1QOR; X-ray; A/B/C/D=177-314.
 DR SMR; P00779; 38-398.
 DR InterPro; IPR004093; Staphylokinase.
 DR InterPro; IPR008124; Streptokinase.
 DR Pfam; PF02821; Staphylokinase; 3.
 DR PRINTS; PR01753; STREPKINASE.
 DR 3D-structure; Direct protein sequencing; Plasminogen activation;
 KW Signal, Virulence.
 FT SIGNAL 1 26
 FT CHAIN 1 26
 FT VARIANT 195 195
 FT VARIANT 207 207
 FT CONFLICT 298 300
 FT CONFLICT 438 438
 FT STRAND 180 180
 FT STRAND 184 194
 FT TURN 199 200
 FT TURN 203 204
 FT STRAND 205 205
 FT STRAND 209 214
 FT TURN 216 217
 FT STRAND 219 221
 FT HELIX 222 236
 FT TURN 238 239
 FT STRAND 240 252
 FT TURN 253 254
 FT STRAND 259 260
 FT STRAND 268 270
 FT STRAND 275 275
 FT STRAND 278 280
 FT TURN 282 284
 FT STRAND 287 289
 FT STRAND 292 304
 FT TURN 305 306
 SQ SEQUENCE 440 AA; 50140 MW; 8FC1F22648AC77A CRC64;

Alignment Scores:

Pred. No.: 2.86e-132 Length: 440
 Score: 1984.00 Matches: 383
 Percent Similarity: 100.00% Conservative: 0
 Best local Similarity: 100.00% Mismatches: 0
 Query Match: 53.25% Indels: 0
 DB: 1 Gaps: 0

US-09-940-235-12 (1-2096) x STRP_STREQ (1-440)

QY 588 ATTGCTGACCTGAGTGGCTGTAGACGTCATCTGTCAACAAGCCAAATTGTTT 647
 DB 27 TLeaAglyProGILtrPLeuLeuAspArgProSeValAsnAsnSerGlnLeuVal 46
 QY 648 AGCGTGTGTACTGTTGAGGGGAGCAATCAAGATTAGTCTTAAATTTTGAATC 707
 DB 47 SerValAlaGlyThrValGlnGlyThrAsnGlnAspIleSerLeuysPheGluIle 66
 QY 708 GATCTAATCATCAGCAGCTGCTCATGAGGAGAAAGACAGAGCAGCTTAAGTCCAATCA 767
 DB 67 AspLeuThrSerArgProAlaHisGlyGlyLysThrGlnGlnGlyLeuSerProLysSer 86
 QY 768 AAACCATTTGCTACGATGATGAGCGGAGATGTCACATTAACCTTGAGAAAGCTGACTTA 827
 DB 87 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlnLysAlaAspLeu 106
 QY 828 AAGCTATTCAAGAACATTTGATCGCTACGTCACAGTAAGCAGCAGTACTTGAAGTC 887
 DB 107 LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 126

QY 888 ATTGATTTTTCAGACGATGCAACCATTAAGTATGATGAAACGCAAGGTCTACTTGTGTAC 947
 DB 127 TLeaSPheAlaSerAspAlaThrIleThrAspArgAsnLysValTyrPheAlaAsp 146
 QY 948 AAAGATGTTTCGGTAACTTTGCCGACCAACCTGTCCAGAAATTTTGTAAAGCGACAT 1007
 DB 147 LysAspGlySerValThrLeuProThrGlnProValGlnGlnPheLeuLeuSerGlyHis 166
 QY 1008 GTCCGCGTTTGACCATTTAAAGAAAAACCAATCAAAACCAAGCGAAATCTGTGATGTG 1067
 DB 167 ValArgValArgProTyrLysGlnLysProIleGlnAsnGlnAlaLysSerValAspVal 186
 QY 1068 GAATATATCTGTACAGTTTATCTCCCTTAACCCGATGACGATTTCAACAGCGTCCAAA 1127
 DB 187 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 206
 QY 1128 GATATTAAGCTATTTGAAAAACATTAAGCTATGCGTGACACATCAATCTCAAGATTACTA 1187
 DB 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeu 226
 QY 1188 GCTCAAGCACAAAGCATTTTAAACAAAAACCCAGGCTATACGATTTATGAACGTGAC 1247
 DB 227 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 246
 QY 1248 TCCTCAATGCTGACATGATGACATTTTCGGTACGATTTTACCAATGATGATCAAGAG 1307
 DB 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGln 266
 QY 1308 TTTACTTACCGTGTAAATTCGGGAAACAGCTTATAGATCAATAAATAATCGTGTCTG 1367
 DB 267 PheThrTyrArgValLysAsnArgGlnGlnAlaTyrArgIleAsnLysLysSerGlyLeu 286
 QY 1368 AATGAAGAAATTAACAACACTGACCTGATGCTTGAGAAATTTACGCTTTAAAAAGG 1427
 DB 287 AsnGlnGlnIleAsnAsnThrAspLeuIleSerGlnLysTyrTyrValLeuLysLysGly 306
 QY 1428 GAAAGCCGTATGATCCCTTGTATGCGACGTCACCTTGAAACGTTCAACATCAATACGTT 1487
 DB 307 GlnLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326
 QY 1488 GATGTGATPACCAACGAATTTGTAAAAAGTGAACAGCTCTTAAACGCTTAACGTAACGTAAC 1547
 DB 327 AspValAspThrAsnGlnGlnLeuLeuLysSerGlnGlnLeuThrAlaSerGlnLysArgAsn 346
 QY 1548 TTAGACTTCAAGATTTTAACGATCTCTGTGATAGGCTTAACCTTCAACAATCTTC 1607
 DB 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
 QY 1608 GATGCTTTGGTATTATGACTATACCTTAACTGGAAGAGTAAAGATTAATCAAGATGAC 1667
 DB 367 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGlnAspAsnHisAspAsp 386
 QY 1668 ACCAACCGTATCAATACCGTTTATATGGCGAAGCAGCCGAGAGAGAGAAATGCTAGTAC 1727
 DB 387 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGlnGlnAsnAlaSerTyr 406
 QY 1728 CATTAGCT 1736
 DB 407 HisLeuAla 409
 RESULT 2
 Q53284_STREQ
 ID Q53284_STREQ PRELIMINARY; PRT; 414 AA.
 AC Q53284;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, last annotation update)
 DE SKC-2.
 OS Streptococcus equisimilis.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=119602;
 RN [1]

RP NUCLEOTIDE SEQUENCE.
 RA Estrada M.P., Hernandez L., Perez A., Rodriguez P., Serrano R.,
 RA Rubiera R., Pedraza A., Padron G., Antuch W., la Fuente J., et al.,
 RT "High level expression of streptokinase in *Escherichia coli*.";
 RL Biotechnology 0:1138-1142 (1992).
 DR EMBL: S66536; AAC60418.1; -; Genomic_DNA.
 DR FDB: 1C4P; X-ray; A/B/C/D=149-285.
 DR SMR: Q53284; 12-372.
 DR GO: GO:0008243; F:plasmaeinnogen activator activity; IEA.
 DR GO: GO:0005515; F:protein binding; IEA.
 DR InterPro: IPR004093; Staphylokinase.
 DR InterPro: IPR008124; Streptokinase.
 DR Pfam: PF02821; Staphylokinase; 3.
 DR PRINTS: PR01753; STREPKINASE.
 DR SEQUENCE 414 AA; 47254 MW; F75BE5831B766904 CRC64;

Alignment Scores:

Pred. No.:	3,256-130	Length:	414
Score:	1955.00	Matches:	377
Percent Similarity:	98.96%	Conservative:	2
Best Local Similarity:	98.43%	Mismatches:	4
Query Match:	52.47%	Indels:	0
DB:	2	Gaps:	0

US-09-940-235-12 (1-2096) x Q53284_STREQ (1-414)

QY 588 ATTGCTGACCTGAGTGGCTGACCGTCCATCTGTCAACAACGCAATGTTGTT 647
 Db 1 MetAlaGlyProGluThrLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
 QY 648 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAATTTTGAATC 707
 Db 21 SerValAlaGlyThrValGlnGlyThrAsnGlnAspIleSerLeuysPheGluIle 40
 QY 708 GATTCAACATCAACGACCTGCTCAAGAGAAAGACAGACGAGCTTAAGTCCAAATCA 767
 Db 41 AspLeuThrSerArgProAlaHisGlyGlyThrGlnGlnGlyLeuSerProLysSer 60
 QY 768 AAACGATTTGCTACTGATGTGGCGGATGTCACATTAATCTTGAGAAAGCTTACTA 827
 Db 61 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGlyValAspLeu 80
 QY 828 AAGCTATTTCAAGAACAAATTTGATCGCTAACGTCACAGTACGACGCTTGAAGTC 887
 Db 81 LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspLysPheGluVal 100
 QY 888 ATTGATTTTGCAGCGATGCAACCATTAAGTCAAGAGGAGGCTTACTTGGCTGAC 947
 Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyValTyrPheAlaAsp 120
 QY 948 AAAGATGTTCCGTTAACTTGGCCGACCCCACTGTCGCAAGATTTTGTACCGGACAT 1007
 Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnGlnPheLeuSerGlyHis 140
 QY 1008 GTGGCGCTTTAGACATATTAAGAAAAACAATTAACAACGAGCAATCTGTGATGTC 1067
 Db 141 ValAlaGlnAlaArgProLysGlyLysProIleGlnAsnGlnAlaLysSerValAspVal 160
 QY 1068 GAATATCTGTGACGTTTACTCCCTTAACCTGATGAGATTTTGACAGAGCTCAAA 1127
 Db 161 GluThrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
 QY 1128 GATTAAGTAACTATTAAGAAACACTAGCTATCGGTGACACCATCATCTCAAGATTA 1187
 Db 181 AspThrLysLeuLeuThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeuLeu 200
 QY 1188 GCTCAAGACCAAGATTTTAACAACAAACCAACGAGCTATACGATTTATGAAGCTGAC 1247
 Db 201 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp 220
 QY 1248 TCCTCAATCGTCACTCATGACATGACATTTCCGTAAGATTTTACCATGATGATCAAGG 1307
 Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProPheAspGlnGlu 240

QY 1308 TTACTTACCGTGTAAAAATCGGAAACAGCTTATAGATCAATRAAAAAATCTGCTGC 1367
 Db 241 PheThrThrIleValLysValAsnArgGlnAlaTyrGluIleAsnLysSerGlyLeu 260
 QY 1368 AATGAGAAATTAACAACACTGACCTGATCTCTGAGAAATATTAAGTCTTAACAAAGG 1427
 Db 261 AsnGlnGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysGly 280
 QY 1428 GAAAAAGCGTATGATCCCTTTGATCGCAGTCACTTGAACGTTTCAATCAATAGCTT 1487
 Db 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
 QY 1488 GATGTCGATCAACCAAGAAATTTGCTAAAAAGTAGACGCTTAACGCTTAGGGAACGTAC 1547
 Db 301 AspValAsnThrAsnGlnLeuLeuLysSerGlnGlnLeuThrAlaSerGluArgAsn 320
 QY 1548 TTAGACTTCAGAGATTTATACGATCTCTGATTAAGGCTAAAGCTTACTTACAAATCTC 1607
 Db 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLysThrAsnLeu 340
 QY 1608 GATGCTTTTGGTATTATGACTATACCTTAACGAAAAAGTAGAGATTAATCAGATGAC 1667
 Db 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyValGluAspAsnHisAspAsp 360
 QY 1668 ACCAAGCGTATCAATACCGTTTATATGGGCAAGGACCCGAAAGAGAAATGCTAGCTAC 1727
 Db 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGlnGlyLysAlaSerLysTyr 380
 QY 1728 CATTTAGCT 1736
 Db 381 HisLeuAla 383

RESULT 3

Q6UK57_STREQ Q6UK57_STREQ PRELIMINARY; PRT; 436 AA.

ID Q6UK57; AC Q6UK57;

DT 05-JUL-2004 (TrEMBLrel. 27, Created)

DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)

DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

DE Streptokinase.

OS Streptococcus equisimilis.

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

OC Streptococcus.

OX NCBI_TaxID=119602;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=ATCC 35666;

RA Costa C.S., Torres F.A.G., Filho S.A.;

RL Submitted (Aug-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL: AY68335; AAQ73571.1; -; Genomic_DNA.

DR SMR: Q6UK57; 59-394.

DR GO: GO:0016301; F:kinase activity; IEA.

DR GO: GO:0008243; F:plasmaeinnogen activator activity; IEA.

DR GO: GO:0005515; F:protein binding; IEA.

DR InterPro: IPR004093; Staphylokinase.

DR InterPro: IPR008124; Streptokinase.

DR Pfam: PF02821; Staphylokinase; 3.

DR PRINTS: PR01753; STREPKINASE.

KW Kinase.

SQ SEQUENCE 436 AA; 49658 MW; CA2DD95F957D8F3C CRC64;

Alignment Scores:

Pred. No.:	2,346-129	Length:	436
Score:	1943.00 <td>Matches:</td> <td>374 </td>	Matches:	374
Percent Similarity:	98.43% <td>Conservative:</td> <td>3 </td>	Conservative:	3
Best Local Similarity:	97.65% <td>Mismatches:</td> <td>6 </td>	Mismatches:	6
Query Match:	52.15% <td>Indels:</td> <td>0 </td>	Indels:	0
DB:	2	Gaps:	0

US-09-940-235-12 (1-2096) x Q6UK57_STREQ (1-436)

QY 588 ATTGCTGACCTGAGTGGCTGTAACCGTCCATCTGTCAACAACGCAATGTTGTT 647

```

Db      23  IleaIaGIProGIuTrIpleuLeuSerAArgProSerValAsnAsnSerGIuLeuValVal 42
Qy      648  AGCGTTGCGGTACTGTTGAGGGGAGCAATCAAGATTAAGCTTAAATTTTGAATC 707
Db      43  SerValaIaGIuThValGIuGIuThraSngInAspIleSerIeuIysPhePheGIuIle 62
Qy      708  GATCTAACATCAGACCTGCTCATGAGAGAAAGACAGACAGGCTTAAGTCCAAATCA 767
Db      63  AspleuThrSerGIuProIaHISGIuGIuGIuThGIuGIuGIuLeuSerProIysSer 82
Qy      768  AAACCATTTGCTACTGATAGTGGCGGAGATGTCACATTAATCTTGAGAAAGCTTA 827
Db      83  LysProPheAlaThAspSerGIuAlaMetProHISLysLeuGIuLysAlaAspLeu 102
Qy      828  AAGCTATTCAGAGAACATTAATGATCGTAAAGCTTCAAGATTAAGACATCTTGAAG 887
Db      103  LysAlaIleGIuGIuGIuGIuLeuIleAlaAsnValHISerAsnAspAspTyPheGIuVal 122
Qy      888  ATTGATTTTGAAGCGATCAACCATTAATCGAATCGAAGGCTTACTTGCTGAC 947
Db      123  IleAspPheAlaSerAspAlaThrIleThraAspAsnGIuLysValTyPheAlaAsp 142
Qy      948  AAAGATGGTTCGGTAACCTTGCCGACCAACCTGTCGAAGATTTTGTAGCGGACAT 1007
Db      143  LysAspGIuSerValThreIeuProThrGIuProValGIuGIuPheLeuLeuGIuHIS 162
Qy      1008  GTGCGCGTTAGACCATTAATTAAGAAAAACCAATTAACCAAGCAATCTGTTAGTGG 1067
Db      163  ValArgValArgProGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIu 182
Qy      1068  GAATATCTAGTACGATTACTCCCTTAACCCGATAGAGATTGAGACAGGCTTCAAA 1127
Db      183  GIuTyTrIuValGIuPheThrProLeuAsnProAspAspPheArgProGIuLeuLys 202
Qy      1128  GATACTAAGCTAATTAAGAAACACTAGCTATCGGTGACACCATCAATCTTAAGATTA 1187
Db      203  AspThrLysIleuLeuLysThreIeuAlaIleGIuAspThrIleThSerGIuGIuLeu 222
Qy      1188  GCTCAAGCAAAAGCATTTTAAACAAACCAACCCGCTATACGATTATTAAGACGTAC 1247
Db      223  AlaGIuAlaGIuInserIleLeuAsnLysThrHISProGIuTyTrIleTyGIuAsp 242
Qy      1248  TCCCTAATCGTACCTCAATGACATGATCCGATGACATTTTCCATTTTACATGATCAAG 1307
Db      243  SerSerIleValThHISAspLysAspIlePheArgThIleIeuProAspGIuGIu 262
Qy      1308  TTTACTTACCGTGTAAATAATCGGAAACAAGCTTAATAGGATCAATTAATAATCGT 1367
Db      263  PheThrTyHISValIleAsnAsnArgGIuGIuAlaTyGIuIleAsnLysLysSerGIuLeu 282
Qy      1368  AATGAAGAATTAACAACAAGCTGATCTGTAGAAAATTAATTCGTTTAAAAAGG 1427
Db      283  AsnGIuGIuIleAsnAsnThraAspLeuIleSerGIuLysTyTrValIleuLysGIu 302
Qy      1428  GAAAGCCGATGATACCTTTGATGGCAGTCACTTGAACCTGTTCCATCAATACGTT 1487
Db      303  GIuLysProIyAspProPheAspArgSerHISLeuLysLeuPheThrIleLysTyVal 322
Qy      1488  GATGCGATTAACAAGAAATTCGTTAAAGAGTGAAGCTTAAACAGCTTACGCAAGTAC 1547
Db      323  AspValaIeThraSngIleuLeuLysSerGIuGIuGIuGIuGIuGIuGIuGIuGIu 342
Qy      1548  TTAGACTTCAGAGATTATACGATCTCGTGAATAGGCTAAATCTACTTAACAATCTTC 1607
Db      343  LeuAspPheArgAspLeuTyAspProArgAspLysAlaLysLeuLeuTyHISAsnLeu 362
Qy      1608  GATGTTTGGATTAATGACATTAATCTTAATCGAAAGATGAGATTAATCAAGTAC 1667
Db      363  AspAlaPheGIuLysIleMetAspTyTrThreIeuThrGIuLysValGIuAsnAsnHISAspAsp 382
Qy      1668  ACCAACCGTATCATTAACGTTATATAGGCAAGCGACCGAAGAGAGATGCTACTAC 1727

```

```

Db      383  ThrAsnArgIleIleThraValTyTrMetGIuLysArgProGIuGIuAsnAlaSerTyR 402
Qy      1728  CATTAGCT 1736
Db      403  HIsLeuAla 405

RESULT 4
STRP_STRS1
ID STRP_STRS1 STANDARD; PRT: 440 AA.
AC P10519;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Streptokinase G precursor.
GN Name=skg;
OS Streptococcus sp. (strain 19909).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=69017;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RX MEDLINE=89160265; PubMed=2922269;
RA Walter F., Siegel M., Malke H.;
RT "Nucleotide sequence of the streptokinase gene from a group-G
RT Streptococcus."
RL Nucleic Acids Res. 17:1262-1262(1989).
CC -!- FUNCTION: This protein is not a protease, but it activates
CC plasminogen by complexing with it. As a potential virulence
CC factor, it is thought to prevent the formation of effective fibrin
CC barriers around the site of infection, thereby contributing to the
CC invasiveness of the cells.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; X13400; CAJ31766.1; -; Genomic DNA.
DR PIR; S02723; S02723.
DR HSSP; P00779; 100R.
DR SMR; P10519; 63-398.
DR InterPro; IPR004093; Staphylokinase.
DR InterPro; IPR008124; Streptokinase.
DR Pfam; PF02821; Staphylokinase; 3.
DR PRINTS; PR01753; STREPKINASE.
KM Plasminogen activation; Signal; Virulence.
FT SIGNAL 1
FT CHAIN 27
FT FT 440 Streptokinase G.
SQ SEQUENCE 440 AA; 50199 MW; 5521F8825FE1B6EA CRC64;

Alignment Scores:
Pred. No.: 2,76e-129 Length: 440
Score: 1942.00 Matches: 375
Percent Similarity: 98.43% Conservative: 2
Best Local Similarity: 97.91% Mismatches: 6
Query Match: 52.12% Indels: 0
DB: 1 Gaps: 0

US-09-940-235-12 (1-2096) x STRP_STRS1 (1-440)
Qy      588  ATTGCTGAGACTGAGTGGCTGTGACCGCTCATCTGTCAACAACAGCAATGGTGT 647
Db      27  IleaIaGIProGIuTrIpleuLeuAspArgProSerValAsnAsnSerGIuLeuValVal 46
Qy      648  AGCGTTGCTGTAATCTGTGAGGGGAGCAATCAAGATTAAGCTTAAATTTTGAATC 707
Db      47  SerValaIaGIuThraValGIuGIuThraSngInAspIleSerIeuLysPhePheGIuIle 66
Qy      708  GATCTAACATCAGACCTGCTCATGAGAGAAAGACAGACAGGCTTAAGTCCAAATCA 767
Db      67  AspleuThrSerGIuProIaHISGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIu 86

```

```

QY 768 AAACATTGCTAGTATGCGCGGAGTGCATTAACCTTGAGAAAGCTGACTACTA 827
Db 87 LysleuPhealathraspergilyAlameProhiblyseuGluysAlaAspleu 106
QY 828 AAGGCTATTCAAGAAATTTGATGCGTAACTGCAAGCTTAAGCTTCTTGAGGTC 887
Db 107 LysAlaIleGlnGlnIleuIleAlaAsnValHisSerAsnAspArglyTyPheGluVal 126
QY 888 ATTGATTTTGAAGCGATGCAACCATTAAGTATGAGAAAGCGAAGTCTCTTGCTGAC 947
Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyValIlyPheAlaAsp 146
QY 948 AAAGATGTTGCGTAACTTTCGCGAACCACTGTCGCAAGATTTTGTCTAAGCGGACAT 1007
Db 147 LysAspGlySerValThrleuProIleGlnProValGlnGlnPheleuLeuLysGlyHis 166
QY 1008 GTGCGCGTTAGACCATTAATAAGAAAAACAATACCAAAACCAAGCGAAATCTGTGATGTG 1067
Db 167 ValArgValArgProlyTyLysGlyLysProValGlnAsnGlnAlaLysSerValAspVal 186
QY 1068 GAATATAGTGTACAGTTTACTCCCTTAAACCTGATGAGCATTTGACAGCGAGTGTCAA 1127
Db 187 GlnlyTyThrValGlnPheThrProleuAsnProAspAspPheArgProAlaLeuLys 206
QY 1128 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACTCTCAAGATTTACTA 1187
Db 207 AspThrLysleuLeuLysThrleuAlaIleGlyAspThrIleThrSerGlnGlnleuLeu 226
QY 1188 GCTCAAGCAAGCAAGCTTTTAAACAAAAACCAACCGGCTATACGATTATGAAAGCTGAC 1247
Db 227 AlaGlnAlaGlnSerIleleuAsnLysAsnHisProGlyTyThrIleTyGlnLysArgAsp 246
QY 1248 TCCCTAATGTCACATGACATGACATGATTCCTGATGCGTAACTTTTACCAAGATCAAGAG 1307
Db 247 SerSerIleValAlaThrHisAspAsnAspIlePheArgThrIleleuProMetAspGlnGln 266
QY 1308 TTTACTTACCGTGTAAATAATCGGGAACAAGCTTATAGATGCAATAAATAATCTGCTGTG 1367
Db 267 PheThrTyThrHisValIleAsnAsnThrAspleuIleSerGlnLysTyTyValIleuLysGly 286
QY 1368 AATGAAAGAAATTAACAACAACATGACCTGATCTCTGAGAAATATTATTCGCTTAAAGAGG 1427
Db 287 AsnGlnGlnIleAsnAsnThrAspleuIleSerGlnLysTyTyValIleuLysGly 306
QY 1448 GAAAGCGGTATGATCCCTTGTATGCGGATGCACTTGAACCTGTCACATCAATATCGTT 1487
Db 307 GlnLysProTyAspProPheAspAspSerHisleuLysleuPheThrIleTyTyVal 326
QY 1488 GATGTCGATACCAAGCAATTTGCTAAAGAGCAGCTCTTAACAGCTAGCGAAAGCTAAC 1547
Db 327 AspValAsnThrAsnGlnleuLeuLysSerGlnGlnleuLeuThrAlaSerGlnLysArgAsn 346
QY 1548 TTAGACTTCAGAGATTTATATGATCCTCGTATAGAGCTTAACTACTCTCAACAATATCTC 1607
Db 347 LeuAspPheArgAspleuTyAspProArgAspLysAlaLysleuLeuTyThrAsnleu 366
QY 1608 GATGCTTTTGTATATGACATTAACCTTAACCTTGAAGAAAGTAAAGAGATTAACGATGAC 1667
Db 367 AspAlaIleGlnIleuIleuSerAspTyThrIleuThrGlyLysValGlnAspAsnHisAspAsp 386
QY 1668 ACCAACCGTATGATACCGTTTATATGAGCAAGCGAACCGGAGAGAGAAATGCTACTATAC 1727
Db 387 ThrAsnAspGlnIleThrValTyMetGlyLysArgProGlnGlnGlnAsnAlaSerTy 406
QY 1728 CATTTAGCT 1736
Db 407 HisLeuAla 409

```

RESULT 5
Q7X0Y1 STRPY
Q7X0Y1 STRPY PRELIMINARY; PRT; 440 AA.
AC Q7X0Y1_01-OCT-2003 (Tremblrel. 25, Created)

```

DT 01-OCT-2003 (Tremblrel. 25, last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, last annotation update)
DE Streptokinase.
CN Name=aka:
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OK NCBI_TaxID=1314;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=D488;
RA PubMed=14679231; DOI=10.1128/JB.186.1.110-121.2004;
RX Kalia A., Bessen D.E.;
RT "Natural selection and evolution of streptococcal virulence genes
RT involved in tissue-specific adaptations.";
RL J. Bacteriol. 186:110-121(2004).
EMBL; AY234137; AAP3957.1; -; Genomic_DNA.
DR HSSP; Q53284; 1C4P.
DR SMR; Q7X0Y1; 63-398.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008243; F:plasmaingn activator activity; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR InterPro; IPR004093; Staphylokinase.
DR Pfam; PF02821; Staphylokinase; 3.
DR PRINTS; PR01753; STREPKINASE.
KW kinase.
SQ SEQUENCE 440 AA; 49903 MW; CC4BEB9647043BAC CRC64;

Alignment Scores:
Pred. No.: 3,43e-120 Length: 440
Score: 1814.00 Matches: 347
Percent Similarity: 95.04% Conservative: 17
Best Local Similarity: 90.60% Mismatches: 19
Query Match: 48.68% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-12 (1-2096) x Q7X0Y1_STRPY (1-440)
QY 588 ATTGCTGACCTGATGAGTGGCTGTAGACCGTTCATCTGTCAACAACAGCAATGTTGTT 647
Db 27 IleAlaGlyTyGlyTyPleuProAspArgProIleAsnAsnSerGlnleuVal 46
QY 648 AGCGTTCGTGTACTGTGTTGAGGGGACGAATCAAGACATTAGCTTAAATTTTGAATC 707
Db 47 SerMetAlaGlyIleValGlnGlyTyThrAspLysValPheIleAsnPhePheGlnIle 66
QY 708 GATCTAACATCAAGACCTGCTCATGAGGAAAGCAGAGCAAGGCTTAAGTCCAAATCA 767
Db 67 AspleuThrSerGlnProAlaHisGlyGlyLysThrGlnGlnIleuSerProLysSer 86
QY 768 AAACATTGCTACTGATAGTGGCGCATGTCACTAACTTGAGAAAGCTGACTACTA 827
Db 87 LysPhePheAlaThrAspAsnGlyAlaMetProhiblyseuGluysAlaAspleu 106
QY 828 AAGGCTATTCAAGAAATTTGATGCGTAACTGCAAGCTTAAAGCTTCTTGAGGTC 887
Db 107 LysAlaIleGlnGlnIleuIleAlaAsnValHisSerAsnAspArglyTyPheGluVal 126
QY 888 ATTGATTTTGAAGCGATGCAACCATTAAGTATGAGAAAGCGAAGTCTCTTGCTGAC 947
Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyValIlyPheAlaAsp 146
QY 948 AAAGATGTTGCGTAACTTTCGCGAACCACTGTCGCAAGATTTTGTCTAAGCGGACAT 1007
Db 147 LysAspGlySerValThrleuProThrGlnProValGlnGlnPheleuLeuLysGlyHis 166
QY 1008 GTGCGCGTTAGACCATTAATAAGAAAAACAATACCAAAACCAAGCGAAATCTGTGATGTG 1067
Db 167 ValArgValArgProlyTyLysGlyLysProValGlnAsnGlnAlaLysSerValAspVal 186
QY 1068 GAATATAGTGTACAGTTTACTCCCTTAAACCTGATGAGCATTTGACAGCGAGTGTCAA 1127

```

Db 187 LysfYrThrValGlnPheThrProLeuAnProAspAspPheArgProGlyLeuLys 206
 QY 1128 GATACGATGCTATGTAACCACTAGCTATCGGTGACACCATCATCTCAAGATTACTA 1187
 Db 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 226
 QY 1188 GCTCAAGCAAAAGCATTTTAAACAAAACCAACCCAGGCTATACGATTTATGAACGTGAC 1247
 Db 227 AlaGlnAlaGlnSerIleLeuAnLysThrHisProGlyTyrThrIleTyrGlnArgAsp 246
 QY 1248 TCCCTCAATCGTCATGACATGACATGATTTTCCGACGATTTTACCATGATGACAAAG 1307
 Db 247 SerSerIleValThrHisAspAnaspIlePheArgThrIleLeuProMetAspGlnGlu 266
 QY 1308 TTTACTTACCGTGTAAATAATGGGAGCAAGCTTTATAGGATCAATTAATAATGGCTG 1367
 Db 267 PheThrTyrArgValLysAspArgGlnGlnAlaTyrGlyIleAnLysLysSerGlyLeu 286
 QY 1368 AATGAAGAATAAACAACATGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGAG 1427
 Db 287 AsnGlnGluIleAsnAsnThrAspLeuIleSerGlnLysTyrIleLeuLysGly 306
 QY 1428 GAAAGCCGTATGATCCCTTTGATGCGATCACTGAACTGTTCACCATCAATACGTT 1487
 Db 307 GluSerProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326
 QY 1488 GATGTCGATACCAAGATTTGCTAAAGTAGAGCGCTCTTAACACTGACGCAAGCTAAC 1547
 Db 327 AspValAsnThrAsnGlnLeuLeuLysSerGlnGlnLeuThrAlaSerGlnAlaGln 346
 QY 1548 TTAGACTTCAGAGATTTATACGATCCCTGATTAAGGCTAAACTACTTACCAACATCTC 1607
 Db 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuThrAsnLeu 366
 QY 1608 GATGCTTTGCTATTTATGACTATACCTTAACGAAAAGTAGAGATTAATCAAGTAGAC 1667
 Db 367 AspAlaPheAspIleMetAspTyrThrLeuThrGlyValGlnAspAnHisAspLys 386
 QY 1668 ACCAAGCTATCATACCGTTTATATATGGGCAAGCAACCCGAAAGAGAGATGCTACTAC 1727
 Db 387 AsnAsnAlaGlnIleValThrValTyrMetGlyLysArgProLysGlyAlaLysGlySerTyr 406
 QY 1728 CATTTAGCT 1736
 Db 407 HisLeuAla 409
 RESULT 6
 Q7X0Y8_STRPY PRELIMINARY; PRT; 440 AA.
 ID Q7X0Y8_STRPY PRELIMINARY; PRT; 440 AA.
 AC Q7X0Y8;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Streptokinase.
 GN Name=eka;
 OS Streptococcus pyogenes.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1314;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=88-019;
 RA PubMed=14679231; DOI=10.1128/DB.186.1.110-121.2004;
 RA Kalita A., Bessen D.E.;
 RT "Natural selection and evolution of streptococcal virulence genes
 RT involved in tissue-specific adaptations.";
 RL J. Bacteriol. 186:110-121(2004).
 DR EMBL; AY234129; AAP39949.1; -; Genomic_DNA.
 DR HSSP; Q53284; 1C4P.
 DR SMK; Q7X0Y8; 63-398.
 DR GO; GO:0016301; F:kinase activity; IEA.
 DR GO; GO:0008243; F:plasminogen activator activity; IEA.
 DR GO; GO:0005515; F:protein binding; IEA.

DR InterPro; IPR004093; Staphylokinase.
 DR InterPro; IPR008124; Streptokinase.
 DR Pfam; PF02821; Staphylokinase; 3.
 DR PRINTS; PR01753; STREPKINASE.
 KW kinase.
 SQ SEQUENCE 440 AA; 49885 MW; 69DB44F026B3975 CRC64;
 Alignment Scores:
 Pred. No.: 2,44e-119 Length: 440
 Score: 1802.00 Matches: 347
 Percent Similarity: 94.52% Conservative: 15
 Best Local Similarity: 90.60% Mismatches: 21
 Query Match: 48.36% Indels: 0
 DB: 2 Gaps: 0
 US-09-940-235-12 (1-2096) x Q7X0Y8_STRPY (1-440)
 QY 588 ATTGCTGACCTTAGAGTGGCTGTAGACCGTCACTGTGTCAACAACACCAATTGGTTGT 647
 Db 27 IleAlaGlyTyrGlyTyrPleuProAspArgProValAsnSerGlnLeuVal 46
 QY 648 AGCGTTGCTGTACTGTGGGGGAGCAATCAAGACTTGTAAATTTTGAATC 707
 Db 47 SerMetAlaGlyIleValGlnGlyThrAspLysValPheIleAsnPheGluIle 66
 QY 708 GATCTAACATCAACGACCTGCTCATGAGGAGAAAGACAGAGCAAGGCTTAAGTCAAAATCA 767
 Db 67 AspLeuThrSerGlnHisAlaHisGlyLysThrGlnGlnLeuSerProLysSer 86
 QY 768 AAACATTTGCTACTGATAGTGGCGCATGTCACATTAACCTTGAGAAAGCTTACTA 827
 Db 87 LysProPheAlaThrAspAsnGlyAlaMetProHisLysLeuGlnLysValAspLeu 106
 QY 828 AAGGCTATTGAAGAACTTGTATCGCTTAAGTCCACGTACACGACTCTTGTAGAGTC 887
 Db 107 LysAlaIleGlnLysGlnLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGlnVal 126
 QY 888 ATTGATTTTGCAGCGCATGCAACCATTAATGATGAAAGGCAAGCTCTTGTGCTGAC 947
 Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnLysValTyrPheAlaAsp 146
 QY 948 AAAGATGCTTCGCTTAACCTTGGCCGACCAACCTGTCACAGAAATTTTGTAAAGCGACAT 1007
 Db 147 LysAspGlySerValThrLeuProthrinGlnProValGlnIleLeuLeuSerGlyHis 166
 QY 1008 GTCCGCGTTAGACCATTAAGAAAGAAACCAATACAAACCAAGCAAAATCTGTGATGTG 1067
 Db 167 ValArgValArgProTyrLysGlnLysProValGlnAsnGlnAlaLysSerValAspVal 186
 QY 1068 GAATATACGTATAGATTTTACCGCTTAACCCGATGACGATTTTCAACGAGCTGAAA 1127
 Db 187 LysfYrThrValGlnPheThrProLeuAnProAspAspPheArgProGlyLeuLys 206
 QY 1128 GATACGATGCTATGTAACCACTAGCTATCGGTGACACCATCATCTCAAGATTACTA 1187
 Db 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 226
 QY 1188 GCTCAAGCAAAAGCATTTTAAACAAAACCAACCCAGGCTATACGATTTATGAACGTGAC 1247
 Db 227 AlaGlnAlaGlnSerIleLeuAnLysThrHisProGlyTyrThrIleTyrGlnArgAsp 246
 QY 1248 TCCCTCAATCGTCATGACATGACATGATTTTCCGACGATTTTACCATGATGACAAAG 1307
 Db 247 SerSerIleValThrHisAspAnaspIlePheArgThrIleLeuProMetAspGlnGlu 266
 QY 1308 TTTACTTACCGTGTAAATAATGGGAGCAAGCTTTATAGGATCAATTAATAATGGCTG 1367
 Db 267 PheThrTyrArgValLysAspArgGlnGlnAlaTyrGlyIleAnLysLysSerGlyLeu 286
 QY 1368 AATGAAGAATAAACAACATGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGAG 1427
 Db 287 AsnGlnGluIleAsnAsnThrAspLeuIleSerGlnLysTyrIleLeuLysGly 306

QY 1428 GAAAAGCCGTATGATCCCTTGTGATCGAGTCACTGTAACCTGTTACCACTCAATACGTT 1487
 |||||
 Db 307 GUSerProtyrAspProPheAspArgSerHisLeuIlySerPheThrIleIlyrVal 326
 |||||
 QY 1488 GATGTCGATACCAACGAATTTGCTAAAGAGACAGCTCTTACAGCTTACGAAACGTAAC 1547
 |||||
 Db 327 AspValAsnThrAsnGlnLeuLeuIlySerGlnGlnLeuThrIleAsnSerGlnArgAsn 346
 |||||
 QY 1548 TTAGACTTCGAGATTTATATACGATCCCTCGATAGAGCTTAACCTACTCTCAACAATCTC 1607
 |||||
 Db 347 LeuAspPheArgAspLeuThrAspProCysAspIlySerIleIlyLeuLeuIlyrAsnAsnLeu 366
 |||||
 QY 1608 GATGCTTTGTGATTATGACATATACCTTACTGTAAGAAAGTAGAGATTAATCAAGATGAC 1667
 |||||
 Db 367 AspAlaPheAspIleMetAspIlyrThrLeuThrGlyIlySerValGlnAspAsnHisAspIly 386
 |||||
 QY 1668 ACCAACCCTATCATACCGCTTTATATAGGACAGCAGCCGACAGAGAGAAATGCTGTTAC 1727
 |||||
 Db 387 AsnAsnArgIleValAlaThrValIlyMetGlyIlyAspProIlySerGlyAlaIlySerGlySer 406
 |||||
 QY 1728 CATTTAGCT 1736
 |||||
 Db 407 HisLeuAla 409
 |||||
 RESULT 7
 Q8KSR8 STRP3 PRELIMINARY; PRT: 440 AA.
 ID Q8KSR8 STRP3 PRELIMINARY; PRT: 440 AA.
 AC Q8KSR8_079W73;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Streptokinase A.
 GN Name=ska; Ordered locus names=SPa1700, SpvM3_1698;
 OS Streptococcus pyogenes (serotype M3).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=301448;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRPAIN-MGAS315 / Serotype M3;
 RX MEDLINE=22133808; PubMed=12122206; DOI=10.1073/pnas.152298499;
 RA Beres S.B., Sylva G.L., Barbican K.D., Lei B., Hoff J.S.,
 RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,
 RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
 RA Schlievert P.M., Musser J.M.,
 RT "Genome sequence of a serotype M3 strain of group A Streptococcus:
 RT phage-encoded toxins, the high-virulence phenotype, and clone
 RT emergence.";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).
 RL [2]
 RN NUCLEOTIDE SEQUENCE.
 RC STRPAIN-SSI-1 / Serotype M3;
 RX MEDLINE=22683278; PubMed=12799345; DOI=10.1101/gr.1096703;
 RA Nakagawa I., Kurikawa K., Yamashita A., Nakata M., Tomiyasu Y.,
 RA Okabeishi N., Kawabata S., Yamazaki K., Shibata T., Yasunaga T.,
 RA Hayashi H., Hattori M., Hamada S.,
 RT "Genome sequence of an M3 strain of Streptococcus pyogenes reveals a
 RT large-scale genomic rearrangement in invasive strains and new insights
 RT into phage evolution.";
 RL Genome Res. 13:1042-1055(2003).
 DR EMBL; AE014169; AAM80305.1; -; Genomic_DNA.
 DR EMBL; BA000034; BAC64795.1; -; Genomic_DNA.
 DR HSSP; P00779; 1L4D.
 DR SMR; Q8KSR8; 63-398.
 DR GO; GO:0016301; F:kinase activity; IEA.
 DR GO; GO:0008243; F:plasminogen activator activity; IEA.
 DR GO; GO:0005515; F:protein binding; IEA.
 DR InterPro; IPR004093; Staphylokinase.
 DR InterPro; IPR008124; Streptokinase.
 DR Pfam; PF02821; Staphylokinase; 3.
 DR PRINTS; PR01753; STREPKINASE.
 KW Complete proteome; kinase.
 SQ SEQUENCE 440 AA; 49911 MW; 7CCBA4F4026B3975 CRC64;

Alignment Scores:
 Pred. No.: 2,446-119 Length: 440
 Score: 1802.00 Matches: 347
 Percent Similarity: 94.52% Conservative: 15
 Best Local Similarity: 94.60% Mismatches: 21
 Query Match: 48.36% Indels: 0
 DB: 2 Gaps: 0
 US-09-940-235-12 (1-2096) x Q8KSR8_STRP3 (1-440)
 QY 588 ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTGCAACAACGCAATGGTGTGTT 647
 |||||
 Db 27 TleAlaGlyrYrGlyrTlPleuProAspArgProProValAsnAsnSerGlnLeuVal 46
 |||||
 QY 648 AGCTTGCTGCTACTGTTTGAGGGGACCAATCAAGCATTAAGCTTTAAATTTTGAATC 707
 |||||
 Db 47 SerMetAlaGlyIleValGlnGlyThrAspIlySerIlyValPheIleAsnPheGlnIle 66
 |||||
 QY 708 GATTAACATCAAGACCGCTCATGAGAGAGACAGGCAAGGCTTAAGTCAAAATCA 767
 |||||
 Db 67 AspLeuThrSerGlnHisAlaHisGlyIlyrThrGlnGlnGlyLeuSerProIlySer 86
 |||||
 QY 768 AAACCATTTGCTACTGATAGTGGCGCATGTCAATTAACCTTGAAGAGCTGACTTACTA 827
 |||||
 Db 87 LysProPheAlaThrAspAsnGlyAlaMetProHisIlySerGlnGlyIlyrAspLeu 106
 |||||
 QY 828 AAGCTATTTCAGACCAATGATGCTTACGTCACAGTAACAGACGACTTGTAGAGTC 887
 |||||
 Db 107 LysAlaIleGlnIlySerGlnLeuIleAlaAsnValHisSerAsnAspGlyrPheGlnVal 126
 |||||
 QY 888 ATTGATTTCGACCGCATGCAACCATTAAGTGAAGACGCAAGCTTACTTGTGCAC 947
 |||||
 Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyIlyrPheAlaAsp 146
 |||||
 QY 948 AAAGATGTTCCGGTAACCTTGCCGACCAACCTGTCCAAGAAATTTTGTAGAGGACAT 1007
 |||||
 Db 147 LysAspGlySerValThrLeuProThrGlnProValGlnGlnPheLeuSerGlyHis 166
 |||||
 QY 1008 GTGGCGGTAGACCATTAATAAGAAACCAATACAAACCAAGCAATCTGTGATGTG 1067
 |||||
 Db 167 ValArgValArgProIlyrIlySerGlnIlyProValGlnAsnGlnAlaIlySerValAspVal 186
 |||||
 QY 1068 GAATTAATCTGTAAGTTACTCTCTTAAACCTGATGACGATTCGACAGGCTCTCAA 1127
 |||||
 Db 187 LysIlyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuIly 206
 |||||
 QY 1128 GATTAAGCTATTGAAACACATGACGATGCGTGACACCAATCATCTCAAGATTTACTA 1187
 |||||
 Db 207 AspThrIlySerLeuLeuIlyrThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeu 226
 |||||
 QY 1188 GCTCAAGCACAAACGATTTTAAACAAACCAACCAAGCTTACGATTTTGAAGCTGAC 1247
 |||||
 Db 227 AlGlnAlaIleGlnSerIleLeuAsnIlyThrHisProGlyrIlyrThrIleIlyrGlnArgAsp 246
 |||||
 QY 1248 TCTCAATGCTCACTCATGACCAATGACATTTTCCGATCGATTTTACCAATGATCAAG 1307
 |||||
 Db 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetArgGlnGln 266
 |||||
 QY 1308 TTACTTACCGGTGTAAATAATCCGGACCAAGCTTATAGATCATTAATAATCTGGCTG 1367
 |||||
 Db 267 PheThrIlyrIlyrValIlyAspArgGlnGlnAlaIlyrGlyIleAsnIlyIlySerGlyLeu 286
 |||||
 QY 1368 AATGAAGAATAAACAACGACGCTGATCTGAGAAATATTAAGTCTTAATAAAAGGG 1427
 |||||
 Db 287 AsnGlnGlnIleAsnAsnThrAspLeuIleSerGlyIlyrIlyrIleLeuIlySerGly 306
 |||||
 QY 1428 GAAAAGCCGTATGATCCCTTGTGATCGAGTCACTGTAACCTGTTACCACTCAATACGTT 1487
 |||||
 Db 307 GUSerProtyrAspProPheAspArgSerHisLeuIlySerPheThrIleIlyrVal 326
 |||||
 QY 1488 GATGTCGATACCAACGAATTTGCTAAAGAGACAGCTCTTACAGCTTACGAAACGTAAC 1547
 |||||

Db	327	AspValaIaenthThraSngLIueneuLeuysSerGIuInleuLeuThrlASerGIuArgSn	346
Qy	1548	TTAGACTTCACAGATTATATACATCTCTGATTAAGGCTAAACTACTTACACAATCTC	1607
Db	347	LeuAspPheArgAspLeuTyrAspProCysAspIysAlaIaLysLeuLeuTyrAsnAsnLeu	366
Qy	1608	GATGCTTTGGTATATAGACATATACCTTAATCGAAAAAGTAGAGATATACAGATAC	1667
Db	367	AspAlaPheAspIleMetAspIyrThrLeuThrGIyIysValGIuAspAsnHisAspLys	386
Qy	1668	ACCAACCCATATATACCGCTTATATATGGGCAAGCAGCAGCCAGAGAGAGAAATCTGATAC	1727
Db	387	AsnAsnAlaGIIleValThrValTyrMetGIyIysArgProIysGIyAlaIaLysGIySerTyr	406
Qy	1728	CATTAGCT	1736
Db	407	HisIleuAla	409
RESULT 8			
ID	STRP_STRPY	STANDARD;	PRT; 440 AA.
AC	P10520;		
DT	01-JUL-1989 (Rel. 11, Created)		
DT	16-OCT-2001 (Rel. 40, Last sequence update)		
DT	10-MAY-2005 (Rel. 47, Last annotation update)		
DE	Streptokinase A precursor.		
GN	Name=sKa; OrderedLocustNames=SPY1979;		
OS	Streptococcus pyogenes.		
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;		
OC	Streptococcus.		
OX	NCBI_TaxID=1314;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RC	STRAIN=SF130/13 / Serotype M1;		
RX	MEDLINE=89160264; PubMed=2646590;		
RA	Walter F. Siegel M., Malke H.;		
RT	"Nucleotide sequence of the streptokinase gene from a Streptococcus		
RL	pyogenes type 1 strain.";		
RL	Nucleic Acids Res. 17:1261-1261(1989).		
RN	[2]		
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].		
RC	STRAIN=SP370 / ATCC 700294 / Serotype M1.		
RX	MEDLINE=21192884; PubMed=11296296; DOI=10.1073/pnas.071559398;		
RA	Ferretti U.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,		
RA	Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,		
RA	Qian Y., Jia H.G., Najjar F.Z., Ren O., Zhu H., Song L., White J.,		
RA	Yuan X., Clifton S.W., Roe B.A., McLaughlin R.E.;		
RL	"Complete genome sequence of an M1 strain of Streptococcus pyogenes.";		
RL	Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).		
CC	-1- FUNCTION: This protein is not a protease, but it activates		
CC	plasminogen by complexing with it. As a potential virulence		
CC	factor, it is thought to prevent the formation of effective fibrin		
CC	barriers around the site of infection, thereby contributing to the		
CC	invasiveness of the cells.		
CC	-----		
CC	This Swiss-Prot entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use as long as its content is in no way modified and this statement is not		
CC	removed.		
CC	-----		
CC	EMBL; X13399; CAA31765.1; -; Genomic DNA.		
DR	EMBL; AE006620; AAK34665.1; -; Genomic DNA.		
DR	PIR; S02724; S02724.		
DR	HSSP; O53284; 1C4P.		
DR	SMR; P10520; 63-398.		
DR	InterPro; IPR004093; Strephylkinase.		
DR	InterPro; IPR008124; Streptokinase.		
DR	Pfam; PF02821; Strephylkinase. 3.		
DR	PRINTS; PR01753; STREPKINASE.		
KW	Complete proteome; Plasminogen activation; Signal; Virulence.		
FT	SIGNAL		
FT	CHAIN	1 26 440	Streptokinase A.

FT	CONFLICT	163	163	L -> V (in Ref. 1).
FT	CONFLICT	345	345	R -> G (in Ref. 1).
FT	CONFLICT	373	373	D -> N (in Ref. 1).
FT	CONFLICT	428	428	D -> Y (in Ref. 1).
FT	CONFLICT	438	438	K -> N (in Ref. 1).
SQL	SEQUENCE	440 AA;	49924 MM;	D6227BF040B758DB CRC64;

Alignment Scores:			
Pred. No.:	9,056-119	Length:	440
Score:	1794.00	Matches:	341
Percent Similarity:	93.99%	Conservative:	19
Best Local Similarity:	89.03%	Mismatches:	23
Query Match:	48.15%	Indels:	0
DB:	1	Gaps:	0

US-09-940-235-12 (1-2096) x STRP_STRPY (1-440)

QY	588	ATTGTGGACCTGAGTGTGCTGTAGACCGTGCATCTGTCAACAAACAGCAATTTGGTGT	647
DB	27	IIleaaglytrgyltrpLeuProasprrgrrProProlleasnSeriGlnleuVal	46
QY	648	AGCGTGTGTGATCTGTGGAGGGAGCAATCAAGACATTAGTCTTAAATTTTGAATC	707
DB	47	SermetlaGlylleValIGluGlyThrAsprrSlySValPheIIeaSnPheGlnIle	66
QY	708	GATCTAAATCAACGACCTGTCTATGAGAGAAAGACAGCAAGGCTTAAATGTCAAATCA	767
DB	67	AspleutrrSerGlnProalHisIleGlyIlyStrGlnGlnGlyLeuSerProIySser	86
QY	768	AAACCATTTGCACTGATGATGAGCGCGCATGTCACATTAACCTTGAGAAAGCTGACTTACTA	827
DB	87	LysProPhealIarrrAspAsnGlyAlaIleMetProHslySleuGlnIlySValIleAspLeuIeu	106
QY	828	AAGGCTATTCAAGAACAATTGATGCTTAACGTCACAGTAAACGACGACTCTTTGAGGTC	887
DB	107	LysAlaIleGlnIlyGlnIleuIIeAlaSnValHisSerAsnAspGlyTYrrPheGlnVal	126
QY	888	ATTGATTTTGGACGGAGTGAACCACTTATCTGATGAAACGGACAGCTTACTTTGCTGAC	947
DB	127	IIeaSpPhealIaSerAspAlaIarrrIleThrAspAgaGlnGlyIlySValTYrrPheAlaAsp	146
QY	948	AAAGATGTTGGATCACTTCGCGCAGCCACCACTGTCCAAAGATTTTGTCTAAGCGGACAT	1007
DB	147	LysAspGlySserValThrLeuProHrrGlnProValGlnGlnPheLeuIeuIlySlyHis	166
QY	1008	GTGCGCGTTAGACCATATTAAGAAACCAATACAAACCAAGCAATCTGTTGATGTG	1067
DB	167	ValaGValaIaGProIyGlyGlnIlySProValGlnaGlnAlaIlySserValaSVal	186
QY	1068	GAATTAATCTGTAACGTTTACTCCCTTAACCTGATGACATTTTGACGACAGGTCGAA	1127
DB	187	GlnTYrrrrValGlnPheThrProIeuSnProAspAspAspPheIarrrProGlyIleuIyS	206
QY	1128	GATACCTAAGCTAATTTGAAAAACATAGCTATCGGTGACACATCACTCAAGAAATTA	1187
DB	207	AsprrrrIlySleuLeuIlyThrIleuAlaIleGlyAsprrrrIleThrSerGlnIlyIleuIeu	226
QY	1188	GCTCAGACCAAGATTTTAAACAAAACCAACCGGCTATACATTTTATGACAGTGC	1244
DB	227	AlaGlnAlaGlnSeriIleIeuAsnIlyStrHisrrProGlyTYrrrrIleTYrrIlyIlyrrGAsp	246
QY	1248	TGCTCAATCGTCACTCATGACAAATGATTTTCCATGATTTTACCAATGGATCAAGAG	1307
DB	247	SerSeriIleValThrHisAspAsnAsprrrrIlePheIlyGlnrrIleIeuProIeChaspGlnIly	266
QY	1308	TTTACTTACCGTGTAAAAATCGGAAACAAGCTTATAGATCAATTAATAATCTGGTGTG	1367
DB	267	PheThrTYrrrrIlySValIlyAsnrrGlnIlyAlaTYrrGlnIlyIleAsnProIyrrThrGlyIle	286
QY	1368	AATGAAGAAATTAACACATGACCTGATCTCTGAGAAATATTAAGTCTTAAAAAAGG	1422
DB	287	LysGlnIlyStrrrAsnrrrrAsprrrrValSeriIlySlyTYrrValIleuIySleuIlyGlnIly	306

QY	1428	GAAGAGCCGTAATGATCCCTTGGATGGAGCTACATTGAAACGTTACACATCAATAGCTT	1487
Db	307	GLIYSPFPIRYTASRPPIHSPRPHSPRGSEHISLEUYSLEUPHETHILIELYTYTVAL	326
QY	1488	GATGTGATATACCAACGAATTGCTATAAAAGTAGAGAGCTTTAAGAGCTAGCGAAGCTAAC	1547
Db	327	ASPVLAIAETHIRASNGLULEULEUYSERGIUGLINLEULEUTHIRLASETCIDLRGASH	346
QY	1548	TTAGACTTCAGAGATTATTAGATCTCTGTGATAGAGCTTAAGCTTACCAACATCTC	1607
Db	347	LEUPSPHEAAGASPLEUTRYASPRCAIRGASPRYSALATYSLEULEUTRYASASINLEU	366
QY	1608	GATGCTTTGGATTATAGSAGCTATACCTTAACCTGGAAAAGTAGAGAAATACAGATAC	1667
Db	367	ASPAALAPHEAPRILEMECASPRTYTRLEUTHRGILYSLVALIGLIDASPANHIAAPRYLS	386
QY	1668	ACCAACCGTATCATATACCGTTTATATAGGGCAAGGACCCGAGAGAGAAATGCTAGTAC	1727
Db	387	ASNDENARGVALVALTHTRVALTYRMETCILIYSHARGPROLYSGILALALYSGLYSERTY	406
QY	1728	CATTTAGCT 1736	
Db	407	HISLEUALA 409	

RESULT 9

Q53ZX6_STRPY PRELIMINARY; PRT; 440 AA.

AC Q53ZX6;

AC 13-SEP-2005 (TREMBLrel. 31, Created)

DT 13-SEP-2005 (TREMBLrel. 31, Last sequence update)

DT 13-SEP-2005 (TREMBLrel. 31, Last annotation update)

DE Streptococcinase.

GN Name=Bka;

OS Streptococcus pyogenes.

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

OC Streptococcus.

OX NCBI_TaxID=1314;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=86-779;

RC PubMed=14679231; DOI=10.1128/JB.186.1.110-121.2004;

RA Kalia A., Bessen D.E.;

RT "Natural selection and evolution of streptococcal virulence genes involved in tissue-specific adaptations.";

RL J. Bacteriol. 186:110-121(2004).

DR EMBL; AY234128; AAP39948.1; -; Genomic_DNA.

KW Kinase.

SO SEQUENCE 440 AA; 49924 MW; D6227BF040B75BDB CRC64;

Alignment Scores:

Pred. No.: 9.05e-119 Length: 440

Score: 1794.00 Matches: 341

Percent Similarity: 93.99% Conservative: 19

Best local Similarity: 89.03% Mismatches: 23

Query Match: 48.15% Indels: 0

DB: 2 Gaps: 0

US-09-940-235-12 (1-2096) x Q53ZX6_STRPY (1-440)

QY 588 ATTGCTGAGCTGAGTGGCTGTAGACGTCATCTGTCAACAACAGCAATGCTTGT 647

Db 27 ILEHAGLYTGYTRPLEPRIOASPRYGRTOIRLEAHENSERGINLEUVAL 46

QY 648 AGCGTGTGCTAGCTGTGAGGGAGCAATCAAGACATTAGCTTAATTTTTGAATC 707

Db 47 SERNEHLAGLYILEVALGUGLYTHRASPRYSLSVALPHEILEANPHEGUILLE 66

QY 708 GATTTAAATCAAGACCTGCTCATATGAGAAAGACAGAGCAAGGCTTAAGTCCAAATCA 767

Db 67 ASPLEUTHRSERGINPROLAHIGLYGYLYETHRGILUGINGLYLEUSERPROLYSER 86

QY 768 AAACATTGTCTACTAGTAGTGGGGCGCATGTCATATAACTGAGAAAGCTGACTTACTA 827

D	b		87	LysProPheAlaIthrAspAsnGlyAlaMetProHisLysLeuGlnLysValAspLeuLeu	106
O	y		828	AAGGTATTCAGAACAATTTGCTGAACGGTCCACAGTAACGACACTATTGAGGTC	887
D	b		107	LysAlaIleGlnLysGlnLeuIleAlaAsnValHisSerAsnAspGlyTyrrPheGluVal	126
O	y		888	ATTGATTTTGCAAGGATGCACACCCTTACTGATCGAAAACGGCAGAGTCTACTTCTGCAC	947
D	b		127	IleAspPheValSerAspAlaThrIleThrAspArgAsnGlyLysValTyrrPheAlaAsp	146
O	y		948	AAAGATGTTGGGTAACTTGCCGCCAACCTGTCCGAAGAATTTTGTCTAAGCGGACAT	1007
D	b		147	LysAspGlySerValThrLeuProThrGlnProValGlnGlnPheLeuLeuLysGlyHis	166
O	y		1008	GTCGCGCTTAGACCATATATAAGAAAAACAATACAAAACAAGCGAAATCTGTGATGTG	1067
D	b		167	ValArgValArgProGlyTyrGlnLysProValGlnAsnGlnAlaLysSerValAspVal	186
O	y		1068	GAATPATCTGTACAGTTTACTCTCTTAAACCTGTATGACGATTTTGACACAGGTCTCAA	1127
D	b		187	GluTyrrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys	206
O	y		1128	GATACATAAGCTATTGAAAAACATGAGCTACGGTGCACACATCACATCCAGAAATTACTA	1188
D	b		207	AspHrtylSLeuLeuLysThrLeuAlaIleGlyAspHrtylIeHisSerGlnLeuLeu	226
O	y		1188	GCTCAGACACAAGCATTTTAAACAAAACAACCCAGGCTATACGATTTATGAACGTGAC	1247
D	b		227	AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrrHrtylGlnLysArg	246
O	y		1248	TCTTCATCTGCATCTCATGACAAATGACATTTTCCGTACGATTTTACCAATGATCAGAG	1307
D	b		247	SerSerIleValIleThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGln	266
O	y		1308	TTTTCTTAACCGTGTAAAAATCGGGAAACAACCTATATGAGCATATAAAAAATCGTGTCG	1367
D	b		267	PheHrtylYrhisValLysAsnArgGlnGlnAlaTyrglnIleAsnProLysThrGlyIle	286
O	y		1368	AATGAGAAATATMAACAACACTGACCTGTACTCTGAGAATATTAACGTCTTAAAAAGGG	1427
D	b		287	LysGlnLysThrAsnAsnThrAspLeuValSerGlnLysTyrrGlyValLeuLysGlnGly	306
O	y		1428	GA AAAACCGGTATGATCCCTTTGATGCGACGTCACTTGAACSTGTCCACCATTAATGCTT	1487
D	b		307	GluLysProTyrrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrrVal	326
O	y		1488	GATGTCGGATACCAACGAATTCGCTAAAAAGTAGAGAGCTTAAACGATGAGGAACGTGAC	1547
D	b		327	AspValAsnThrAsnGlnLeuLysLysSerGlnGlnLeuLeuThrAlaSerGlnLysArg	346
O	y		1548	TTAGACTTCAGAGATTATATACGATCTCTGNTAGTAAGGCTAAACTACTTCAACAATCTC	1607
D	b		347	LeuAspPheArgAspLeuTyrrAspProArgAspLysAlaLysLeuLeuTyrrAspAsnLeu	366
O	y		1608	GATGCTTTTGGTATTATGACTATACCTTAACTGAAAAGTAGAGATATCACGATGAC	1667
D	b		367	AspAlaPheAspIleMetAspTyrrThrLeuThrGlyLysValGlnAspAsnHisAspLys	386
O	y		1668	ACCAACCGTATCANAAACGTTTATATGGGCAAGGACCCGAAGAGAGAAAGCTAGTAC	1727
D	b		387	AsnAsnArgValValThrValTyrrMetGlyLysArgProLysGlyAlaLysGlySerTyrr	406
O	y		1728	CATTAGCT 1736	
D	b		407	HisLeuAla 409	
<hr/>					
RESULT 10					
ID	Q7XOY2_STRPY	PRELIMINARY;	PRT;	440 AA.	
ID	Q7XOY2_STRPY	PRELIMINARY;	PRT;	440 AA.	
DT	01-OCT-2003	(Tremblrel. 25, Created)			
DT	01-OCT-2003	(Tremblrel. 25, Last sequence update)			
DT	01-MAR-2004	(Tremblrel. 26, Last annotation update)			

DE Streptokinase.
 GN Name=ska;
 OS Streptococcus pyogenes.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 NC NCBI_TaxID=1314;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=J306;
 RX PubMed=14679231; DOI=10.1128/JB.186.1.110-121.2004;
 RA Kalia A., Besen D.B.;
 RT "Natural selection and evolution of streptococcal virulence genes
 involved in tissue-specific adaptations.";
 RL J. Bacteriol. 186:110-121(2004).
 DR EMBL: AY234136; AAP39956.1; -; Genomic_DNA.
 DR HSSP: O53284; 1CAP.
 DR SMK: Q7X072; 63-398.
 DR GO: GO:0016301; F:kinase activity; IEA.
 DR GO: GO:0008243; F:plasmaenogen activator activity; IEA.
 DR GO: GO:0005515; F:protein binding; IEA.
 DR InterPro: IPR004093; Staphylokinase.
 DR InterPro: IPR008124; Streptokinase.
 DR Pfam: PF02821; Staphylokinase; 3.
 DR PRINTS: PRO1753; STRBPKINASE.
 KW Kinase.
 SQ SEQUENCE 440 AA; 50188 MW; 4CB42AC2A9062C2E CRC64;

Alignment Scores:

Pred. No.: 2,846-118 Length: 440
 Score: 1787.00 Matches: 342
 Percent Similarity: 94.26% Conservative: 19
 Best Local Similarity: 89.30% Mismatches: 22
 Query Match: 47.96% Indels: 0
 DB: 2 Gaps: 0

US-09-940-235-12 (1-2096) x Q7X0Y2_STRPY (1-440)

QY 588 ATGTGGAGCCGTGAGTGGTGTGACCGTCCATCTGTCAGACAGCAATTTGTTT 647
 Db 27 TLeaaaglyrYrGlyrTprleuprOasphapropProlleasnsersetInleuVal 46
 QY 648 AGCGTGTGCTGACTGTGTGAGGAGGAGATCAAGCATTTAGTCTTAAATTTTGAATC 707
 Db 47 SeretelaaglylleValGluclythrasplysValphelIeasnphegeIuile 66
 QY 708 GATCTACATCAACGACCTGCTCATGAGAGAAAGACAGAGCAAGCTTAAAGTCCAAATCA 767
 Db 67 AspleuthrserGlnProIahisglYglysthrGlnGlnGlyleuserProIyser 86
 QY 768 AAACCATTTGCTACGTAGTAGGCGGAGATGTCATTAACCTTGAGAAAGCTGACTTA 827
 Db 87 LysProphelaathrAsphasnsersalmetProhIslyleuglnlysaIaAspleu 106
 QY 828 AAGCTATTCAAGAACATTTGATCGCTAAAGTCCACAGATGAGAGACTTGTAGGTC 887
 Db 107 LyslaaIaIeGlnlyGlnleuIlelaaenValhIsersAspGlyrYrhegluVal 126
 QY 888 ATTGATTTTGAAGGATGCAACCATTTACTGATGAAAGCGGAGGCTTACTTTGCTGAC 947
 Db 127 TLeaspphealaaserAspAlaathrIleThrasparGasnGlyValrYrphelaAsp 146
 QY 948 AAAAGATTTGGGTAACTTGTCCGACCAAGCTGTCACAGAAATTTTGTAGCGGACAT 1007
 Db 147 LysaspGlyserValThrleuprOthrGlnProValGlnGlnPhleuIeuThrGlyhIs 166
 QY 1008 GTGCGCGTTAGACCATATAAGAAAAACAATACAAACCAAGCGAAATCTGTGATGTC 1067
 Db 167 ValArgValaIasgProIyrlYsGlylsProValGlnasGlnAlaIaserservalAspVal 186
 QY 1068 GAATATATCTGTACAGTTTACTCCCTTAAACCTGATGAGATTTGACAGCAAGCTCAAA 1127
 Db 187 GluYrThrValGlnPhethrProleuasnProaAspAspPhelysProValleuYs 206

QY 1128 GATACATAAGCTATTGAAACAACACTAGTATCGTGACACATCAATCTCAAGATTACTA 1187
 Db 207 AspThrlylsleuIeuYslylsleuAlaIeGlyAspThrValThrsersGlnleuLeu 226
 QY 1188 GCTCAAGCACAAAGCATTTTAAACAAAACCAACCGGCTTAAAGATTATGAAAGTGC 1247
 Db 227 AlaGlnlaIeGlnserIleleuasnGlnserHisProaspYrThrIleYrGluArgAsp 246
 QY 1248 TCGCTAATGCTCATCTATGACATATGACATTTTCCGTACATTTTACCAAGATGACAG 1307
 Db 247 SerSerIleValThrIhsasplysAspIlephargThrIleuPrometAspGlnGlu 266
 QY 1308 TTACTTACCGTGTGTTAAATTCGGGAAACAAGCTTATGATCATATAAAATCGTGTG 1367
 Db 267 PheThrTythrGlyValIysasnaIasgGlnAlaIstYrGlyIleasnYslysserGlyLeu 286
 QY 1368 AATGAAGAAATTAACAACACTGACCTGATCTGAGAAATATTACGCTTAAAGAGG 1427
 Db 287 AsnGlnGlnIleasnThrAspleuIleSerGlnYsYrYrValleuYslyeArg 306
 QY 1428 GAAAGCCCGTATGATCCCTTGTATCGGACGTCATTTGAAACTGTTCACATCAATAGTT 1487
 Db 307 GluYsPProIyrlAspProphAspAspSerHisleuYsleuPheThrIleYsYrVal 326
 QY 1488 GATGTGATPCCAAAGCAATGTCTTAAAGAGGAGAGCTTAAACAGCTTACGAGCTAAC 1547
 Db 327 AspValasnThrleuGlnleuYslysserGlnGlnleuIeuThraIasergIuYhAsn 346
 QY 1548 TTAGACTTCAGAGATTATACATCTCGTATAGGCTTAAAGCTTCAACATCTC 1607
 Db 347 LeuAspPheArgAspleuYrAspProArgAspLyalaYsleuIeuYrAsnaasnleu 366
 QY 1608 GATGCTTTGTGATTATGACATATCTTACTGAGAAAGATAGATATATCAGATGAC 1667
 Db 367 AspAlaPheaspIlewetAspYrThrleuThrGlylsValGlnAspAsnHisAspLy 386
 QY 1668 ACCAAGCTATCAACCGTTTATATGGGCAAGCAAGCAAGAGAGAAATGCTAGCTAC 1727
 Db 387 AsnaasnArgValaIaThrValYrMetGlylsYsArgProYsGlyAlaYsGlyserYr 406
 QY 1728 CATTTAGCT 1736
 Db 407 HisleuAla 409

RESULT 11
 ID Q7X0Y7_STRPY PRELIMINARY; PRT; 440 AA.
 AC Q7X0Y7;
 DT 01-OCT-2003 (T-EMBLrel. 25, Created)
 DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
 DE Streptokinase.
 GN Name=ska;
 OS Streptococcus pyogenes.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 NC NCBI_TaxID=1314;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=IRP112;
 RX PubMed=14679231; DOI=10.1128/JB.186.1.110-121.2004;
 RA Kalia A., Besen D.B.;
 RT "Natural selection and evolution of streptococcal virulence genes
 involved in tissue-specific adaptations.";
 RL J. Bacteriol. 186:110-121(2004).
 DR EMBL: AY234130; AAP39950.1; -; Genomic_DNA.
 DR HSSP: O53284; 1CAP.
 DR SMK: Q7X0Y7; 63-398.
 DR GO: GO:0016301; F:kinase activity; IEA.
 DR GO: GO:0008243; F:plasmaenogen activator activity; IEA.
 DR GO: GO:0005515; F:protein binding; IEA.
 DR InterPro: IPR004093; Staphylokinase.
 DR InterPro: IPR008124; Streptokinase.

DR Pfam: PF02821; Staphylokinase; 3.
 DR PRINTS: PR01753; STREPKINASE.
 KW Kinase.
 SQ SEQUENCE 440 AA; 50160 MW; 1CSAF07907EC7AC8 CRC64;

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score: 1758.00	3,276-116	440	338	19	26	0
Best Local Similarity:	93.21%					
Query Match:	47.18%					

US-09-940-235-12 (1-2096) x Q7X0Y7_STRPY (1-440)

```

QY 588 ATTGCTGACCTGAGTGGCTGTAGACCGTCATCTGTCAACAACGCAATTGTTGTT 647
DB 27 lleaiaGlyTYrclYlrpleuenaPArgProProValasnbnSerGlnleuValVal 46
QY 648 AGCGTGTCTGTACTGTGTGAGGAGCAATCAAGACATTAGTCTTAATTTTGAATC 707
DB 47 SerMetAlaGlylValGlnGlyThrAspLysLysValPheIleAsnPhePheGlnIle 66
QY 708 GATCTAACATCACACCTGCTCATGAGAGAAAGACAGAGAGGCTTAACTCCAAATCA 767
DB 67 AspLeuThrSerGlnProAlaHisGlyGlyLysThrGlnGlnGlyLeuSerProLysSer 86
QY 768 AAACCACTTTGCTATCTGATAGTGGCGGAGATGTCATTAATTTGAGAAAGCTTACTA 827
DB 87 LysProPheAlaThrAsnSerSerAlaMetProHisLysLeuGlnLysAlaAspLeu 106
QY 828 AAGGCTATTCAAGAACATTTGATGCTAAGCTCACAGTACAGCACTTCTTGAAGTGC 887
DB 107 LysAlaIleGlnGlnArgLeuIleAlaAsnValHisSerAsnAspGlyTYrPheGlnVal 126
QY 888 ATTGATTTTTCAGAGCGATGCAACATTACTGATGAAACGCGAAGGCTTACTTGTGAC 947
DB 127 lleaPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTYrPheAlaAsp 146
QY 948 AAAGATGCTTGGCTTAACTTGGCGAGCCCAACCTGTCAGAGATTTTGTCAACGGAGAT 1007
DB 147 LysAspAspSerValThrLeuProThrGlnProValGlnGlnLeuLeuArgGlyHis 166
QY 1008 GTGCGGGTTAGACCATATTAAGAAAACAATGACAAACCAAGCGAAATGTTGATGTC 1067
DB 167 ValAlaGlyValArgProTYrLysGlyLysProIleGlnThrProAlaLysSerValAspVal 186
QY 1068 GAATATACTGTACAGTTTACTCCCTTAAACCTGATGACGATTTGACAGCGGCTCAAA 1127
DB 187 ArgTYrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProValIleuLys 206
QY 1128 GATCTAACCTATTGAAAACACTAGCTATGCTGAGACCATTCACATCTCAAGAAATTA 1187
DB 207 AsnThrLysLeuLeuLysThrLeuAlaIleGlyGlyTYrValThrSerGlnGlnLeuLeu 226
QY 1188 GCTCAAGCAACAAGATTTTAAACAAAACCCAGCGCTATTCGATTTTGAACGTGAC 1247
DB 227 AlaGlnAlaGlnSerIleLeuAsnGlnSerHisAspAspTYrThrIleTYrGlnAspAsp 246
QY 1248 TCCCTCAATCGTCACATGACATGACATTTTCCGATGATTTTCAACATGATCAAGAG 1307
DB 247 SerSerIleValThrHisAspAsnAspIlePheAspThrIleLeuProMetAspGlnGln 266
QY 1308 TTACTTACCGCTGTTAAATACGGGAACAAGCTTATAGATCAATTAATAAATCGTCTG 1367
DB 267 PheThrTYrHisIleLysAspArgGlnGlnAlaTYrGlyTYrIleAsnLysLysSerGlyGln 286
QY 1368 AATGAAGAAATTAACAACACTGACTGATCTGAGAAATATTAAGTCTTAAAAAGG 1427
DB 287 GlnGlnLysThrAsnAsnThrAspLeuIleSerGlnLysTYrTYrValIleuLysGly 306
QY 1428 GAAAAGCGGTATGATCCCTTATGTCAGACATTTGAAACTGTTCACCAATCAATCGTT 1487

```

```

DB 307 GlnLysProTYrAspProPheAspArgSerHisLysLysLeuPheThrIleAsnTYrVal 326
QY 1488 GATGTCGATACCAACGCAATTTGCTAAAGAGAGAGCTCTTAAACGCTAGCGAAGCTAAC 1547
DB 327 AspValAsnThrAsnLysLysLeuLysSerGlnGlnLeuLeuThrAlaSerGlnAsn 346
QY 1548 TTAGACTTCAGAGATTTTATACGATCCCTGATTAAGGCTTAAACTCTTCAACATCTC 1607
DB 347 LeuAspPheArgAspLeuTYrAspProArgAspLysAlaLysLeuLeuTYrAsnAsnLeu 366
QY 1608 GATGCTTTTGGTATTTATGACCTATACCTTACGTGAAAGATGAAGATTAATCACGATGAC 1667
DB 367 AspAlaPheGlyIleMetAspTYrThrLeuThrGlyLysValGlnAspAsnHisAspLys 386
QY 1668 ACCAAGCTATCATTAACCGTTTATAGCGGAGAGCGAAGGCAAGGAAATGCTAGTAC 1727
DB 387 AsnAsnArgValValThrValTYrMetGlyLysArgProGlnGlnAlaSerTYr 406
QY 1728 CATTTAGCT 1736
DB 407 HisLeuAla 409

RESULT 12
Q5X9T6_STRP6 PRELIMINARY; PRT; 440 AA.
ID Q5X9T6_STRP6
AC Q5X9T6;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Streptokinase (EC 3.4.-.-).
GN OrderedclonNames=M6_Spy1692;
OS Streptococcus pyogenes (serotype M6).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=301450;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MGA510394;
RX PubMed=15272401; DOI=10.1086/422697;
RA Banks D.J., Porcella S.F., Barbican K.D., Beres S.B., Phillips L.E.,
RA Voylich J.M., Deleo F.R., Martin J.M., Somerville G.A., Musser J.M.;
RT "Progress toward characterization of the group A Streptococcus
RT metagenome: complete genome sequence of a macrolide-resistant serotype
RT M6 strain".
RL J. Infect. Dis. 190:727-738(2004).
DR EMBL; CP000003; AAT87827.1; -; Genomic_DNA.
DR SMR; Q5X9T6; 63-398.
DR GO; GO:0016787; F:Hydrolase activity; IEA.
DR GO; GO:0016301; F:Kinase activity; IEA.
DR GO; GO:0008243; F:Plasminogen activator activity; IEA.
DR GO; GO:0005151; F:Protein binding; IEA.
DR InterPro; IPR004093; Staphylokinase.
DR InterPro; IPR008124; Streptokinase.
DR Pfam; PF02821; Staphylokinase; 3.
DR PRINTS; PR01753; STREPKINASE.
KW Complete proteome; Hydrolase; Kinase.
SQ SEQUENCE 440 AA; 50160 MW; 1CSAF07907EC7AC8 CRC64;

Alignment Scores:
Pred. No.: 3,276-116 Length: 440
Score: 1758.00 Matches: 338
Percent Similarity: 93.21% Conservative: 19
Best Local Similarity: 88.25% Mismatches: 26
Query Match: 47.18% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-12 (1-2096) x Q5X9T6_STRP6 (1-440)
QY 588 ATTGCTGACCTGAGTGGCTGTAGACCGTCATCTGTCAACAACGCAATTGTTGTT 647
DB 27 lleaiaGlyTYrclYlrpleuenaPArgProProValasnbnSerGlnleuValVal 46
QY 648 AGCGTGTCTGTACTGTGTGAGGAGCAATCAAGACATTAGTCTTAAATTTTGAATC 707

```

```

Db      47 SerMetAlaGlyIleValIleGluGlyThrAspLysLysValPheIleAsnPhePheGluIle 66
Qy      708 GATCTAAACATCAAGCACTGCTCTATGAGAAAGACAGACAGCAAGCTTAAGTCCAAATATCA 767
Db      67 AspLeuThrSerGlnProAlaHisGlyGlyLysThrGluGlnIleLysLeuSerProLysSer 86
Qy      768 AAACCATTTGCTACTATGATGCGGAGATGTCACATAACTTGAGAAAGCTGACTTACTA 827
Db      87 LysProPheAlaThrAsnSerSerAlaMetProHisLysLeuGluLysAlaAspLeuLeu 106
Qy      828 AAGGCTATTCAAGAACATTTGATCGCTTAACGTCCACAGTAACGACGACTTACTTTGAGGTC 887
Db      107 LysAlaIleGlnGluArgLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal 126
Qy      888 ATTGATTTTGGCAAGCATGACCATTAATCTGATCGAAAGCGGAGGCTTACTTTGCTGAC 947
Db      127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146
Qy      948 AAAGATGGTTCGGTAACTTGGCCGACCCGACCGCTGCCAAGATTTTGGCTAAGCGGACAT 1007
Db      147 LysAspAspSerValThrLeuProHisGlnProValGlnGluPheLeuLeuArgGlyHis 166
Qy      1008 GTGCGCGTTAGACCATATATAAGAAAAACCAATACAAACCAAGCGAATCTGTTGATGTG 1067
Db      167 ValArgValArgProTyrLysGluLysProIleGlnThrProAlaLysSerValAspVal 186
Qy      1068 GAATTTACTGTAACGTTTACTCCCTTAAACCTGATGACGATTTTCAGACCGAGCTCAAA 1127
Db      187 ArgTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProValLeuLys 206
Qy      1128 GATCTAAGCTATTGAAAACACTAGTACGTGTCACACCATCACTCAAGAAATTAACTA 1187
Db      207 AsnThrLysLeuLeuLysThrLeuAlaIleGlyGlyThrValThrSerGlnGluLeuLeu 226
Qy      1188 GCTCAAGACCAAGAATTTTAAACAAACCAACCCAGGCTTAACGATTTATGAAACGTGAC 1247
Db      227 AlaGlnAlaGlnSerIleLeuAsnGlnLysSerHisProAspTyrThrIleTyrGluAspAsp 246
Qy      1248 TCTCTAATGTCACCTCAGTACCAATGCAATTTCCGTAAGATTTTACCAATGATCAAGAG 1307
Db      247 SerSerIleValThrHisAspAsnAspIlePheAspThrIleLeuProMetAspGlnGlu 266
Qy      1308 TTTACTTACCGCTTAATAATGGGAAACAGCTTATAGATCAATTAATAATTTGGTCTG 1367
Db      267 PheThrTyrHisIleLysAspArgGluGlnAlaTyrGlyIleAsnLysLysSerGlyGln 286
Qy      1368 AATGAAGAAATTAACAACACACTGACCTGATCTCTGAGAAATTTTACGCTTAAAAAAGG 1427
Db      287 GluGluLysThrAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 306
Qy      1428 GAAAAACCGTATGATCCCTTTGATGCGAGTCACTGAAACTGTCCACATCAATACGTT 1487
Db      307 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleAsnTyrVal 326
Qy      1488 GATGTCGATCCAGCAATTTGCTTAATAAAGTGCAGCTCTTAAACAGCTTCCGCAAGCTAAC 1547
Db      327 AspValAsnThrAsnLysLeuLysSerGlnGlnIleLeuThrLaseGluAsnGaa 346
Qy      1548 TTAGACTTCAGAGATTTATAGCATCTCGTGAATAAGGCTAAACTATCTACAACAATCTC 1607
Db      347 LeuAspPheAlaArgAspLeuTyrAspProAlaAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
Qy      1608 GATGCTTTTGTATTATGAGCATTAATCTTAACTGAAAAAGTAGAGATTAATCAAGTAC 1667
Db      367 AspAlaPheGlyIleMetAspTyrThrIleuThrGlyLysValGluAspAsnHisAspLys 386
Qy      1668 ACCAACCGTATCATTAACGTTTATATGGCAAGCGACCCGAAAGAGAGAAATGCTACTAC 1727
Db      387 AsnAsnArgValValThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 406
Qy      1728 CATTAGGCT 1736

```

```

Db      407 HisLeuAla 409
RESULT 13
Q8NZAG_STRP8
ID Q8NZAG_STRP8 PRELIMINARY; PRT; 440 AA.
AC Q8NZAG;
DT 01-OCT-2002 (T-EMBLrel. 22, Created)
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Streptokinase.
GN OrderedLocustNames=spym18_2042;
OS Streptococcus pyogenes (serotype M18) .
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
NCBI_TaxID=301451;
RX MEDLINE=21927593; PubMed=11917108; DOI=10.1073/pnas.062526099;
RA Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
RA Sylva G.L., Sturdevant D.E., Ricklefs S.M., Forcella S.F.,
RA Parkins L.D., Berez S.B., Campbell D.S., Smith T.M., Zhang Q.,
RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
RT "Genome sequence and comparative microarray analysis of serotype M18
RT group A Streptococcus strains associated with acute rheumatic fever
RT outbreaks."
RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673 (2002).
DR EMBL; AE010107; AL98517.1; -; Genomic_DNA.
DR HSSP; O53284; 1C4P.
DR SMR; Q8NZAG; 63-398.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008243; F:plasmaenogen activator activity; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR InterPro; IPR004093; Staphylokinase.
DR InterPro; IPR008124; Streptokinase.
DR Pfam; PF02821; Staphylokinase; 3.
DR PRINTS; PR01753; Streptokinase.
KW Complete proteome; kinase.
SQ SEQUENCE 440 AA; 50382 MW; 47DAB3557EBB88E CRC64;

Alignment Scores:
Pred. No.: 5,34e-116 Length: 440
Score: 1755.00 Matches: 334
Percent Similarity: 93.73% Conservative: 25
Best Local Similarity: 87.21% Mismatches: 24
Query Match: 47.10% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-12 (1-2096) x Q8NZAG_STRP8 (1-440)
Qy      588 ATTGCTGACCTGAGTGGCTGCTAGACCGTCACTGTCACAAACAGCAATGTTGTT 647
Db      27 IleAlaGlyTyrGluThrLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 46
Qy      648 AGCGTTGCTGTAAGTCTTGAAGGGAAGCAATCAAGATTAAGTCTTAAATTTTGAAGC 707
Db      47 SerMetAlaGlyIleValIleGluGlyThrAspLysLysValPheIleAsnPhePheGluIle 66
Qy      708 GATCTAATCAAGCAAGCTGCTCATGAGAGAAAGACAGACAGCAAGCTTAAGTCCAAATATCA 767
Db      67 AspLeuThrSerGlnProAlaHisGlyGlyLysThrGluGlnIleLysLeuSerProLysSer 86
Qy      768 AAACCATTTGCTACTATGATGCGGAGATGTCACATAACTTGAGAAAGCTGACTTACTA 827
Db      87 LysProPheAlaThrAspAsnGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu 106
Qy      828 AAGGCTATTCAAGAACATTTGATCGCTTAACGTCCACAGTAACGACGACTTACTTTGAGGTC 887
Db      107 LysAlaIleGlnGluArgLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal 126
Qy      888 ATTGATTTTGGCAAGCATGACCATTAATCTGATCGAAAGCGGAGGCTTACTTTGCTGAC 947
Db      127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146

```

```

QY 948 AAGATGTTGGTAACTTGGCCGACCACTGTCCAGAAATTTTGTAAAGCGACAT 1007
Db 147 GlnABgLySerValThrLeuProThrGlnProIleGlnGlnPheLeuAlaGlyHis 166
QY 1008 GGGCCCGTTAGACCAATTAAGAAAAACCAATCAAAACCAAGCGAAATCTGTGATGTG 1067
Db 167 ValArgValArgProTyrLysGlyLysProIleGlnThrProAlaLysSerValAspIle 186
QY 1068 GAATATACGTGTACAGTTACTCCCTTAACCCGTAGACGATTTCAGACAGGCTGAAA 1127
Db 187 ArgTyrThrValGlnPheThrProLeuAsnProAspAspPheLysProValLeuLys 206
QY 1128 GATATTAAGCTATTGAAAAACATAGCTATCGGTGACACATCAATCTCAAGATTACTA 1187
Db 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 226
QY 1188 GCTCAAGCACAAGCAATTTTAAACAAAACCAACCCAGGCTATACGATTATGAACGTGAC 1247
Db 227 AlaGlnAlaGlnSerIleLeuGlnGlnSerHisSerAspTyrThrIleTyrGlnArgAsp 246
QY 1248 TCCTCAATGCTACATGACATGACATGATTCCTGATGATTTTACCAATGATCAAGAG 1307
Db 247 SerSerIleValIleThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
QY 1308 TTTACTTACCGTGTAAATAATCGGGAACAAGCTTATAGATCAATAAATAATCTGGTCTG 1367
Db 267 PheThrTyrHisIleLeuAspArgGlnGlnAlaTyrCylIleAsnLysLysSerGlnGln 286
QY 1368 AATGAAGAAATTAACAACAACATGACCTGATCTGTGAGAAATATTACGCTTAAAGAGG 1427
Db 287 GlnGlnLysThrAsnAsnThrAspLeuIleSerGlnLysTyrTyrValLeuLysGly 306
QY 1428 GAAAAGCCCTAGATATCCCTTTGATCGCAGTCACTTGAACCTGTTCACCAATACGTT 1487
Db 307 GlnLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleAsnTyrVal 326
QY 1488 GATGTCGATTAACAACAAGATTGTGTAATAAAGTAGACAGCTCTTAACAGTAGCGAAGCAAC 1547
Db 327 AspValAsnThrAsnLysLeuLeuLysSerGlnGlnLeuLeuThrAlaSerGlnArgAsn 346
QY 1548 TTAAGCTTCAGAGATTATACATCTCTGATTAAGGCTTAACCTACTTACCAACAACTTC 1607
Db 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
QY 1608 GATGCTTTTGGTATTATGACATATACCTTAACCTGAAAGTAGAAGATTAATCAAGATGAC 1667
Db 367 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGlnAspAsnHisAspLys 386
QY 1668 ACCAACCGATACATAACGCTTATATATGGCAAGCGAACCGCAAGAGAAGATGCTAGTAC 1727
Db 387 AsnAsnArgValValThrValTyrMetGlyLysArgProGlnGlnGlnAsnAlaSerTyr 406
QY 1728 CATTAAGCT 1736
Db 407 HisLeuAla 409

```

```

RX PubMed=14679231; DOI=10.1128/JB.186.1.110-121.2004;
RA Kaia A., Beesen D.E.;
RT "Natural selection and evolution of streptococcal virulence genes
RT involved in tissue-specific adaptations.";
RU J. Bacteriol. 186:110-121(2004).
DR EMBL: AY234134; AAF39554.1; -; Genomic_DNA.
DR HSSP; Q53284; 1C4P.
DR SMR; Q7X0Y3; 63-398.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008243; F:plasmaeogen activator activity; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR InterPro; IPR004093; Staphylokinase.
DR Pfam; PF02821; Staphylokinase; 3.
DR PRINTS; PR01753; STREPKINASE.
KW kinase
SQ SEQUENCE 440 AA; 50342 MW; B3C6DBA50DB2C39 CRC64;

Alignment Scores:
Pred. No.: 1e-113 Length: 440
Score: 1723.00 Matches: 327
Percent Similarity: 92.43% Conservative: 27
Best Local Similarity: 85.38% Mismatches: 29
Query Match: 46.24% Indels: 0
DB: Gaps: 2

US-09-940-235-12 (1-2096) x Q7X0Y3_STRPY (1-440)
QY 588 ATTGCTGACCTGAGTGGCTGTGACCGTGCATCTGTCAACCAACGCAATGGTGTG 647
Db 27 IleAlaGlyTyrGlnTyrPheLeuAspArgProSerValAsnAsnSerGlnLeuVal 46
QY 648 AGCGTTGCTGATCTGTGAGGGAGCAATCAAGCATTAATCTTAATTTTGAATC 707
Db 47 SerMetAlaGlyIleValGlnGlyThrAspLysValPheIleAsnPheGlnIle 66
QY 708 GATCAATCAATCAACGACCGCTCATGAGGAGAAAGACAGGCAAGGCTTAAGTCAAAAATCA 767
Db 67 AspLeuThrSerGlnProAlaHisGlyGlyLysThrGlnGlnGlnLysLeuSerProLysSer 86
QY 768 AAACATTTGCTACTGATAGTGGCGCATGTCACTTAACCTTGAAGAGCTGACTTA 827
Db 87 LysProPheAlaThrAspAsnGlyAlaMetProHisLysLeuGlnLysAlaAspLeuLeu 106
QY 828 AAGCTATTGAACAACATTTGATGCTTACGTCACAGTAACGACACTTGTGAGGTC 887
Db 107 LysAlaIleGlnGlnArgLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGlnVal 126
QY 888 ATTGATTTTGCAAGCGATGCAACCATTAAGTGAAGCAAGGCAAGCTACTTGTGAC 947
Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAspAsnIleTyrPheAlaAsn 146
QY 948 AAGATGTTCCGATTAACCTTGGCCGACCACTGTCCAGAAATTTTGTAAAGCGACAT 1007
Db 147 GlnABgLySerValThrLeuProThrGlnProIleGlnGlnPheLeuAlaGlyHis 166
QY 1008 GGGCCCGTTAGACCAATTAAGAAAAACCAATCAAAACCAAGCGAAATCTGTGATGTG 1067
Db 167 ValArgValArgProTyrLysGlyLysProIleGlnThrProAlaLysSerValAspIle 186
QY 1068 GAATATACGTGTACAGTTACTCCCTTAACCCGTAGACGATTTCAGACAGGCTGAAA 1127
Db 187 ArgTyrThrValGlnPheThrProLeuAsnProAspAspPheLysProValLeuLys 206
QY 1128 GATATTAAGCTATTGAAAAACATAGCTATCGGTGACACATCAATCTCAAGATTACTA 1187
Db 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeuLeu 226
QY 1188 GCTCAAGCACAAGCAATTTTAAACAAAACCAACCCAGGCTATACGATTATGAACGTGAC 1247
Db 227 AlaGlnAlaGlnSerIleLeuGlnGlnSerHisSerAspTyrThrIleTyrGlnArgAsp 246
QY 1248 TCCTCAATGCTACATGACATGACATGATTCCTGATGATTTTACCAATGATCAAGAG 1307

```

```

|||||
Db      247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProThrAspGlnGlu 266
OY      1308 TTTACTTACCGGTGTTAAATAATGGGAAACAGCTTATAGATCAATAAATAATCGTCTG 1367
Db      267 PheThrTyHisValIleAspAsnArgGlnGlnAlaTyGlnAsnAspAsnIleThrGlnGlu 286
OY      1368 AATGAAGAATAAACAACACTGACCTGATCTCTGAGAAATATATACCTCTTAAATAAGG 1427
Db      287 LysLysGlnThrLysAsnThrAspLeuIleSerGlnLysTyThrIleLeuLysGln 306
OY      1428 GAAAAGCGGTATGATCCCTTTGATGCGAGTCACTTGAACCTGTTCACATCAATACGTT 1487
Db      307 GluLysProTyAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyVal 326
OY      1488 GATGTCGATACCAAGCAATGCTTAAATAATGAGCAGCTCTTAAACAGCTAGCAAGCTAAC 1547
Db      327 AspValAspThrLysAspLeuLysSerGlnGlnLeuThrAlaSerGlnArgAsn 346
OY      1548 TTAGACTTCAGAGATTTATACGATCTCGTGAATAGGCTTAACCTACTTCAACAATCTC 1607
Db      347 LeuAspPheArgAspLeuTyAspProArgAspLysAlaLysLeuLeuTyAsnAsnLeu 366
OY      1608 GATGCTTTGTTATATGACTATACCTTACTGAGAAAGTAGAGATATACAGATGAC 1667
Db      367 AspAlaPheAspIleMetAspTyThrThrIleThrGlnLysValGlnAspAsnHisAspLys 386
OY      1668 ACCAACCGTATCATTAACGTTTATATGCGCAAGCGACCCGAGAGAGAGAAATGCTACCTAC 1727
Db      387 AsnAsnArgValValThrValTyMetGlnLysArgProLysGlnLysLysGlnSerTy 406
OY      1728 CATTAAGCT 1736
Db      407 HisLeuAla 409

```

RESULT 15
Q7X0Y0_STRPY :PRELIMINARY; PRT; 440 AA.

AC Q7X0Y0_ 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Streptokinase.
GN Name=eka;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus;
OX NCBI_TaxID=1314;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=d633;
RX PubMed=14679231; DOI=10.1128/JB.186.1.110-121.2004;
RA Kalita A., Beesen D.E.;
RT "Natural selection and evolution of streptococcal virulence genes
involved in tissue-specific adaptations.";
RL J. Bacteriol. 186:110-121(2004).
DR EMBL: AY224138; AAP3958.1; -; Genomic_DNA.
DR HSSP: Q53284; ICAP.
DR SMR: Q7X0Y0; 63-398.
DR GO: GO:0016301; F:kinase activity; IEA.
DR GO: GO:0008243; F:plasmaenogen activator activity; IEA.
DR GO: GO:0005515; F:protein binding; IEA.
DR InterPro: IPR004093; Staphylokinase.
DR InterPro: IPR008124; Streptokinase.
DR Pfam: PF02821; Staphylokinase.
DR PRINTS: PR01753; STREPKINASE.
KW Kinase.

SEQUENCE 440 AA; 50143 MW; 4ACB2C9349D680C CRC64;

Alignment Scores: 3.71e-113 Length: 440
Pred. No.: 1715.00 Matches: 325
Score: 91.91% Conservative: 27
Percent Similarity: 91.91%

```

Best Local Similarity: 84.86% Mismatches: 31
Query Match: 46.03% Indels: 0
DB: 2 Gaps: 0
US-09-940-235-12 (1-2096) x Q7X0Y0_STRPY (1-440)
OY      588 ATTGCTGACCTGATGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCAATGGTTGTT 647
Db      27 IleAlaGlnProGlnThrLysLeuGlnGlnArgProValAsnAsnSerGlnLeuVal 46
OY      648 AGCGTGTCTGTACTGTGAGGGAGCAATCAAGACATTAAGTCTTAAATTTTGAATC 707
Db      47 SerMetAlaGlnLysLeuValGlnGlnTyThrAspLysLysValPheIleAsnPheGlnIle 66
OY      708 GATCTAACATCAGACAGCTGCTCATGAGAGAAAGACAGACAGCAAGCTTAAGTCCAAATCA 767
Db      67 AspLeuThrSerGlnProAlaHisGlnGlnLysThrGlnGlnGlnLysSerProLysSer 86
OY      768 AAACCAATTTGCTACTGATAGTGGCGGATGTCAATTAACCTTGAGAAAGCTTACTA 827
Db      87 LysProPheAlaThrAspAsnGlnLysAlaMetProHisLysLeuGlnLysAlaAspLeu 106
OY      828 AAGCTATTCAAGAACCAATTTGATGCTCAACGTCACAGTAAAGCACTACTTTGAGATC 887
Db      107 LysAlaIleGlnGlnLysGlnLeuIleAlaAsnValHisSerAsnAspGlnTyThrPheGlnVal 126
OY      888 ATTGATTTTGCACGCGATGCAACCATTTCTGATCGAAAGCGCAAGCTTACTTGTGAC 947
Db      127 IleAspPheAlaSerAspAlaThrIleThrAspArgAspAsnIleTyThrPheAlaAsn 146
OY      948 AAGATGGTTCGGTAACTTGGACCGACCAACCGTCCAAAGATTTTGTAAAGCGACAT 1007
Db      147 GlnAspGlnSerValThrLeuProThrGlnProIleGlnIlePheLeuLysGlnHis 166
OY      1008 GTGCGGTTAGACCATTAATAAGAAAAACCAATCAAAACCAAGCAATCTGTGATGTG 1067
Db      167 ValArgValArgProTyThrLysGlnLysProIleGlnThrProAlaLysSerValAspIle 186
OY      1068 GAATTAATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTGACACAGGCTCAAA 1127
Db      187 ArgTyThrValGlnPheThrProLeuAspProAspAspPheLysProValIleLys 206
OY      1128 GATTAACGCTATTTGAAAACACTAGCTATCGGAGCAACCATCAATCTCAAGAAATTA 1187
Db      207 AspThrLysLeuLeuLysThrLeuAlaIleGlnLysAsnThrIleThrSerGlnLysLeu 226
OY      1188 GCTCAAGCAACAAGCAATTTTAAACAATAACCAACCGAGCTATACGATTTTGAACGTGAC 1247
Db      227 AlaGlnAlaGlnSerIleLeuAsnGlnSerHisProAspTyThrIleTyThrLysArg 246
OY      1248 TCCTCAATGTCATCTGATGACAAATGACATTTTCGTAACGATTTTACCAATGATCAAGAG 1307
Db      247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProThrAspGlnGlu 266
OY      1308 TTTACTTACCGGTGTTAAATAATGGGAAACAGCTTATAGATCAATAAATAATCGTCTG 1367
Db      267 PheThrTyHisValIleLysAsnArgGlnGlnAlaTyLysLysAlaAsnSerLysThrAspIle 286
OY      1368 AATGAAGAATAAACAACACTGACCTGATCTCTGAGAAATATATACGCTCTTAAATAAGG 1427
Db      287 LysGlnLysThrAsnAsnThrAspLeuIleSerGlnLysTyThrIleLysLysGln 306
OY      1428 GAAAAGCGGTATGATCCCTTTGATGCGAGTCACTTGAACCTGTTCACATCAATACGTT 1487
Db      307 GluLysProTyAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyVal 326
OY      1488 GATGTCGATACCAAGCAATGCTTAAATAATGAGCAGCTCTTAAACAGCTAGCAAGCTAAC 1547
Db      327 AspValAspThrLysAlaLeuLeuLysSerGlnGlnLeuThrAlaSerGlnArgAsn 346
OY      1548 TTAGACTTCAGAGATTTATACGATCTCGTGAATAGGCTTAACCTACTTCAACAATCTC 1607
Db      347 LeuAspPheArgAspLeuTyAspProArgAspLysAlaLysLeuLeuTyAsnAsnLeu 366

```

Qy	1608	GATGCTTTGGTATTATGACTATACCTTA	CTGAAAGTAGAGATAATCAGATGAC	1667
Db	367	AspAlaPheaspIleMetAspTyrThrLeu	ThrGlySerValGluAspAsnHisAspLys	386
Qy	1668	ACCAACCGTATCATACCGTTTATATGGG	CAAGCACCAGAGAGAGAAATGCTAGCTAC	1727
Db	387	AsnAsnArgValValThrValTyrMetGly	LysArgProLysGlyAlaLysGlySerTyr	406
Qy	1728	CATTAGCT	1736	
Db	407	HisLeuAla	409	

Search completed: January 28, 2006, 02:34:30
 Job time : 245.882 secs